AFWAL-TR-81-1142

AD A108607

IMPROVEMENTS AND EXTENSIONS OF THE GEOMETRICAL DILUTION OF PRECISION (GDOP) CONCEPT FOR SELECTING NAVIGATION MEASUREMENTS

William L. Brogan
Department of Electrical Engineering
University of Nebraska-Lincoln

September 1981

Final Report for Period 1 May 1980 to 1 May 1981

Approved for public release; distribution unlimited

AVIONICS LABORATORY
AIR FORCE WRIGHT AERONAUTICAL LABORATORIES
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture use, or sell any patented invention that may in any way be related thereto.

This report has been reviewed by the Office of Public Affairs (ASD/PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

WILLIAM E. SHEPHARD, Tech Mgr
Reference Systems Analysis and
Evaluation Group
Systems Avionics Division
Avionics Laboratory

RONALD L. RINGO
Chief, Reference Systems Branch
Systems Avionics Division
Avionics Laboratory

FOR THE COMMANDER:

RICHARD H. BOIVIN, Colonel, USAF
Chief, System Avionics Division
Avionics Laboratory

"If your address has changed, if you wish to be removed from our mailing list, or if the addressee is no longer employed by your organization please notify AFWAL/AAAN-3, W-P AFB, OH 45433 to help us maintain a current mailing list".

Copies of this report should not be returned unless return is required by security considerations, contractual obligations, or notice on a specific document.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFWAL-TR-81-1142	2. GOVT ACCESSION NO.: AD-A108 607
3. RECIPIENT'S CATALOG NUMBER:	
4. TITLE (and Subtitle): IMPROVEMENTS AND EXTENSIONS OF THE GEOMETRICAL DILUTION OF PRECISION (GDOP) CONCEPT FOR SELECTING NAVIGATION MEASUREMENTS	5. TYPE OF REPORT & PERIOD COVERED: Final Report for Period 1 May 80 - 1 May 81
	6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): William L. Brogan	8. CONTRACT OR GRANT NUMBER(s): USAF F33615-80-C-1092
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Electrical Engineering, University of Nebraska-Lincoln, Lincoln, Nebraska 68588	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: A 30201 6095 14 11
11. CONTROLLING OFFICE NAME AND ADDRESS: Avionics Laboratory (AFWAL/AAAN), Air Force Wright Aeronautical Laboratories (AFSC), Wright-Patterson Air Force Base, Ohio 45433	12. REPORT DATE: September 1981
	13. NUMBER OF PAGES: 72
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

covariance analysis, eigenvalues, eigenvectors, error ellipsoid, Global Positioning Satellite system, Kalman filter, minimum variance, navigation errors, optimal measurement selection, sensor substitution strategy, suboptimal filtering, weighted least-squares.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The optimal measurement selection problem is studied for the Global Positioning Satellite system (GPS). The fundamentals of the Geometrical Dilution of Precision (GDOP) concept are briefly reviewed, because GDOP is frequently used to rank the effectiveness of potential measurements. It is shown that GDOP does not always select the best measurements in the sense of minimum mean square navigation error. For this reason weighted least-squares and minimum variance (Kalman filter) methods are used to derive several

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

401023

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract, continued

improved measurement ranking schemes.

Nonuniform measurement noise and a priori knowledge about the state are taken into account. Any specified weighted combination of the component error variances can be minimized. Eigenvalue-eigenvector theory is used to derive useful bounds and to provide geometrical insight. A purely algorithmic approach is also presented and applied to a number of representative GPS cases.

In addition to the static (single time point) selection problem, the time-sequential selection problem is considered. It is shown that the choice giving the smallest instantaneous position error may not give the smallest time average position error because of the trade-off existing with velocity errors.

A promising approach is formulated for the sensor substitution problem. That is, should degraded (jammed) satellite channels be deleted from the measurements, and if so, when? Which backup sensors should be used instead? A suboptimal filtering approach using state vector partitioning is used to derive some preliminary answers to these questions.

FOREWORD

This report was written by Professor William L. Brogan of the Department of Electrical Engineering at the University of Nebraska-Lincoln. The work was carried out under USAF Contract Number F33615-80-C-1092 for Reference Systems Branch, System Avionics Division, Avionics Laboratory, (AFWAL/AAAN), Air Force Systems Command, Wright-Patterson Air Force Base, Ohio.

The work reported here was carried out during the period 1 May 1980 through 1 May 1981. The author was assisted by Mr. Walter Warnsholz, a graduate student in the Department of Electrical Engineering at the University of Nebraska-Lincoln. The guidance and direction of several people at the Avionics Laboratory was most helpful. This especially includes Mr. William E. Shephard as well as Mr. Stanton H. Musick and Col. Robert Edwards of the Reference Systems Branch. Credit for typing the manuscript is due Mrs. Eunice Everett.

TABLE OF CONTENTS

TABLE OF CONTENTS (concluded)

1. INTRODUCTION

The Geometrical Dilution Of Precision (GDOP) concept is a powerful, simple and widely used measure for evaluating the effectiveness of potential measurements. As its name implies, GDOP characterizes the geometrical effects present when using multiple measurements to obtain a solution to the navigation problem.

A simple illustration is the determination of user position in a plane, from measurements of angles to two known reference points. Position determination reduces to finding the intersection of the two straight lines emanating from the reference points and passing through the user's unknown position. If the lines are nearly parallel, their intersection point is very sensitive to small inaccuracies in measuring angular directions. This is a case of poor GDOP. More robust solutions, i.e., good GDOP, results when the two lines are orthogonal.

As important as geometry is, it is not the only consideration. It will be shown that selecting measurements based solely on geometry does not necessarily produce the optimum reduction of navigation uncertainties. This study re-examines the question of measurement selection. In addition to geometry the varying noise levels on candidate measurements are considered, as well as the a priori navigation state uncertainties. Section 4 considers generalization of the GDOP concept to include these factors.

When a sequence of measurements are made over a period of time, the optimum measurements may not be those which give the most accurate position solution at each point in time. This is true for two different kinds of reasons. First, error growth between measurements is a strong function of velocity errors, so a position/velocity error trade-off exists. Second, instantaneous criteria like GDOP do not account for visibility times, or lost measurement opportunities during search and acquisition of new measurements. The result is that a sequence of less effective measurements over a longer period of time may produce smaller average errors than would the most effective measurements when used over a shorter time.

There is also strong evidence that a law of dimenishing returns applies for sequential measurements. If the best measurement combination (based on instantaneous measures like GDOP) is used repeatedly over a sequence of measurement times, the incremental value decreases with each use. All other things being equal, it is possible that some sort of alternation among several combinations of (instantaneously) less effective measurements will give a smaller average navigation error. Section 5 is devoted to some of the issues related to sequential measurements.

If a primary navigation sensor fails completely, the system is forced to rely on some less optimum mode of operation, perhaps as crude as dead reckoning. However, when a primary navigation system is operational but at a degraded accuracy level, the decision is not so clear. Section 6 discusses some aspects of the sensor substitution problem. In the Global Positioning Satellite system (GPS), measurements from some satellites may be heavily jammed, or lost completely. The question of sensor substitution deals with conditions under which a certain satellite channel ought to be dropped, and if so, should certain states like clock bias stop being updated temporarily, or should some substitute set of inertial or radar, etc. hardware be used to carry on in a degraded mode?

The scope of this study contract, and the specific tasks alluded to above are reiterated in Section 3. Prior to this more formal statement of the problem, Section 2 provides some necessary background development, including a mathematical review of GDOP.

2. BACKGROUND

This section presents a brief tutorial development of GDOP from the point of view of least-squares solutions to simultaneous linear equations. Then weighted least-squares solutions, with and without a priori statistics, are reviewed. Finally, the recursive Kalman filter equations for minimum variance estimation are presented. The well-known equivalence (under certain conditions) of weighted least-squares and minimum variance solutions is pointed out. Comparison of these various results will bring to light certain deficiencies in the GDOP approach, and will suggest generalizations of GDOP, to be developed in Section 3.

2.1 An Introduction to GDOP Using Least-Squares

Consider a vector of measurement components, $\underline{z}$, corrupted by additive noise $\underline{n}$ and linearly related to the state vector $\underline{x}$.

$$\underline{z} = H\underline{x} + \underline{n} \tag{1}$$

An estimate of $\underline{x}$, called $\hat{\underline{x}}$, is sought from the given $\underline{z}$ vector. One approach to this problem is the least-squares estimator $\hat{\underline{x}}$. That is, select $\hat{\underline{x}}$ so as to minimize the sum of the squares of the components of the difference $\underline{z} - H\hat{\underline{x}}$. This means $\hat{\underline{x}}$ should minimize

$$J_1 = (\underline{z} - H\hat{\underline{x}})^T(\underline{z} - H\hat{\underline{x}}) \tag{2}$$

Differentiating J_1 with respect to $\hat{\underline{x}}$ and setting the result to zero gives

$$- H^T\underline{z} + H^TH\hat{\underline{x}} = \underline{0}$$

so that the well-known least-squares estimate is

$$\hat{\underline{x}} = (H^TH)^{-1} H^T\underline{z} \tag{3}$$

The quality of this estimate can be assessed by forming the error

$$\tilde{\underline{x}} = \hat{\underline{x}} - \underline{x} \tag{4}$$

Using Eq. (1) and Eq. (2) gives

$$\begin{aligned}\tilde{\underline{x}} &= (H^TH)^{-1} H^TH\underline{x} + (H^TH)^{-1} H^T\underline{n} - \underline{x} \\ &= (H^TH)^{-1} H^T\underline{n}\end{aligned} \tag{5}$$

If the noise $\underline{n}$ has zero mean, taking the expected value $E\{\}$ of Eq. (5), shows that the estimator error $\tilde{\underline{x}}$ also has a zero mean. The covariance of the estimation error can be related to the covariance of the noise $\underline{n}$.

$$E\{\tilde{\underline{x}}\,\tilde{\underline{x}}^T\} = (H^TH)^{-1}\,H^T\,E\{\underline{n}\,\underline{n}^T\}H(H^TH)^{-1} \tag{6}$$

If all components of $\underline{n}$ are pairwise uncorrelated and have unit variance, i.e.,

$$E\{n_i\,n_j\} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

then $$E\{\underline{n}\,\underline{n}^T\} = I$$

so that

$$E\{\tilde{\underline{x}}\,\tilde{\underline{x}}^T\} = (H^TH)^{-1}\,H^TH(H^TH)^{-1}$$

$$= (H^TH)^{-1} \tag{7}$$

The matrix $(H^TH)^{-1}$ will be called the GDOP matrix. Various scalar terms derived from this matrix are frequently used. For example, the trace of $(H^TH)^{-1}$, written $TR(H^TH)^{-1}$ is the sum of the squares of all error components when $E\{n_i\,n_j\} = \delta_{ij}$. The square root, $[TR(H^TH)^{-1}]^{\frac{1}{2}}$ is usually called GDOP.[1,2] The square root of the sum of all diagonal terms which correspond to position components is sometimes called PDOP.[1,2] If the horizontal plane diagonal components of $(H^TH)^{-1}$ are summed, the square root is called HDOP.[1,2] In this study, major attention is directed to $TR(H^TH)^{-1}$. Omission of the square root will not alter any conclusions regarding the minimization of GDOP.

It is clear that all GDOP-related performance measures indicate the error in an estimated navigation quantity "per unit of measurement noise" covariance. Deviations from this assumption are considered in the next Section.

All of the above GDOP-related measures depend solely on the geometry matrix H. Smaller GDOP values indicate stronger or more robust geometric solutions to the estimation problem. For these reasons, when some freedom exists in the choice of measurements,

good (i.e., small) GDOP is often used as the selection criterion. It will be shown later that this is not always the best choice.

2.2 Weighted Least-Squares Solution

The "per unit noise" concept inherent in GDOP is not so useful (or more bluntly, invalid) when certain measurement components are noisier than others. When a choice exists between two possible sets of measurements, the set with poorer GDOP may be preferable if they are of sufficiently higher accuracy. This trade-off was pointed out in conjunction with an integrated GPS/Inertial Navigation System.[3] There, the desire to achieve good GDOP suggested selection of range measurements separated by angles near 90°. However, this caused an increase in atmospheric diffraction errors in those measurements with low elevation angles. Another potentially more severe problem with nonuniform noise arises in a jamming environment. Some of the highly directional antennas may receive large amounts of noise pollution. This degradation would more than overcome any GDOP advantages.

If the measurement noise covariance is not just the unit matrix, but rather is $E\{\underline{n}\,\underline{n}^T\} = I\sigma^2$ then the GDOP matrix just ratios up by the scalar σ^2. If all potential measurements have this same variance σ^2, the choice of the best measurement set will still be the best GDOP set. The multiplication scalar σ^2 will not affect the relative rankings.

The measurement noise is called nonuniform when different measurements have different noise levels, as indicated by their variances. In this case a weighted least-squares approach[4] to estimation is often used. The quadratic form of Eq. (2) is modified by inserting a weighting matrix W.

$$J_2 = (\underline{z} - H\hat{\underline{x}})^T W(\underline{z} - H\hat{\underline{x}}) \tag{8}$$

It is customary to select the weighting matrix as the inverse of the noise covariance matrix (i.e., $W = R^{-1}$ where $R = E\{\underline{n}\,\underline{n}^T\}$). This choice weights the accurate measurements more, the noisy ones less.

Minimizing this modified cost function leads to a modification of Eq. (2)

$$\hat{\underline{x}} = (H^T R^{-1} H)^{-1} H^T R^{-1} \underline{z} \tag{9}$$

and a corresponding modification of Eq. (6) or (7)

$$E\{\tilde{\underline{x}}\,\tilde{\underline{x}}^T\} = (H^T R^{-1} H)^{-1} \tag{10}$$

Equation (10), and generalizations of it, will be discussed in terms of potential GDOP generalizations in Section 4. It is presented here only to contrast the traditional GDOP with the resulting estimation error covariance.

Thus far no mention has been made of the knowledge about the navigation states which may exist prior to making navigation measurements. In almost all real situations prior knowledge will exist because of previous measurements or for other reasons. The uncertainty volume for $\underline{x}$ need not be spherical. Some components of $\underline{x}$ can have much more uncertainty associated with them, leading to elongated ellipsoidal uncertainty volumes. Those components with the largest uncertainty are said to have "maximum need" for improvement. Some measurements with less than optimal over-all GDOP, may be more effective in reducing navigation uncertainties.

The treatment of a prior information in estimation can be approached in several ways. One way[4] is to again modify Eq. (2), this time to

$$J_3 = (\underline{z} - H\hat{\underline{x}})^T R^{-1} (\underline{z} - H\hat{\underline{x}}) + (\underline{x}_o - \hat{\underline{x}})^T P_o^{-1} (\underline{x}_o - \hat{\underline{x}}) \tag{11}$$

where $\underline{x}_o$ represents an a priori estimate of $\underline{x}$ and the covariance of $\underline{x}_o$ is P_o. Minimizing J_3 amounts to a trade-off between measurement residuals and deviations from $\underline{x}_o$. This leads to a modified estimator

$$\hat{\underline{x}} = [H^T R^{-1} H + P_o^{-1}]^{-1} [H^T R^{-1} \underline{z} + P_o^{-1} \underline{x}_o] \tag{12a}$$

$$= \underline{x}_o + P_o H^T [H P_o H^T + R]^{-1} [z - H\underline{x}_o] \tag{12b}$$

Form (12a) most commonly is used in weighted least-squares analysis, while form (12b) is one of the well-known Kalman filter equations[5].

These two forms are in fact identities[6], and the corresponding forms of the a posteriori estimation error covariance matrix are given by the following identities.[6]

$$E\{\tilde{\underline{x}}\,\tilde{\underline{x}}^T\} = P_1 = [H^T R^{-1} H + P_o^{-1}]^{-1} \tag{13a}$$

$$= P_o - P_o H^T [H P_o H^T + R]^{-1} H P_o \tag{13b}$$

Note that Eq. (13a) looks most like the GDOP matrix in that it reduces to Eq. (10) if $P_o \to \infty$ (very large initial uncertainties) and it reduces to Eq. (7) if, in addition, $R = I$. These equations will be used in Section 3, when discussing a generalized GDOP concept. They are presented here to enable the deficiencies of the traditional GDOP method to be demonstrated.

2.3 An Introductory Example

Consider a planar navigation problem. Range measurements from the unknown vehicle location $\underline{x}$ to two known points $\underline{x}_a$ and $\underline{x}_b$ are used. The geometry is shown in Figure 1. The range measurements are z_a and z_b.

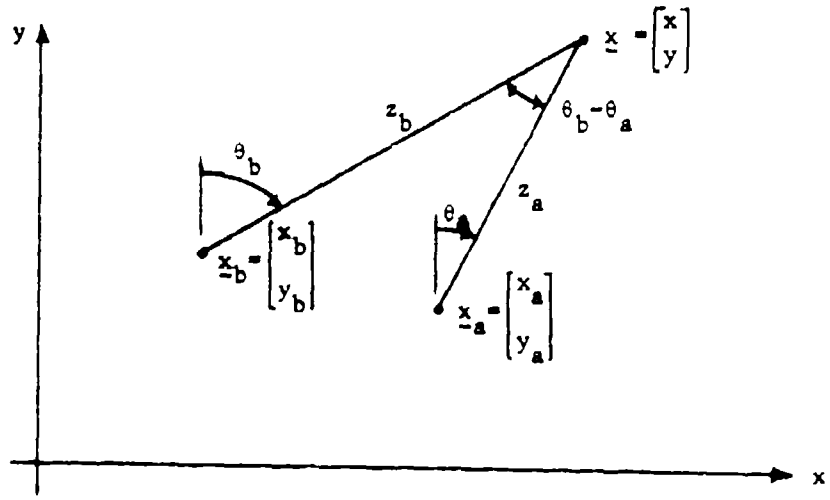

Figure 1. Geometry of a Planar Range Navigation System

Neglecting measurement noise, the ranges are

$$z_a = \sqrt{(x-x_a)^2 + (y-y_a)^2} \tag{14}$$

$$z_b = \sqrt{(x-x_b)^2 + (y-y_b)^2}$$

These form components of the nonlinear measurement equation

$$\underline{z} = \underline{h}(\underline{x}) \tag{15}$$

In order to use GDOP in this example, linear sensitivity equations are required. Linearizing about some nominal point $\underline{x}_n$ gives

$$\Delta\underline{z} = \begin{bmatrix} \frac{\partial z_a}{\partial x}\Delta x + \frac{\partial z_a}{\partial y}\Delta y \\ \frac{\partial z_b}{\partial x}\Delta x + \frac{\partial z_b}{\partial y}\Delta y \end{bmatrix}$$

$$= \begin{bmatrix} \frac{x_n - x_a}{z_a} & \frac{y_n - y_a}{z_a} \\ \frac{x_n - x_b}{z_b} & \frac{y_n - y_b}{z_b} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \tag{16}$$

$$\text{or } \Delta\underline{z} = H\,\Delta\underline{x} \tag{17}$$

The components in H can be written as trig functions by using the geometry of Figure 1.

$$H = \begin{bmatrix} \sin\theta_a & \cos\theta_a \\ \sin\theta_b & \cos\theta_b \end{bmatrix} \tag{18}$$

The GDOP matrix is expressible as

$$(H^T H)^{-1} = \frac{1}{\sin^2(\theta_a - \theta_b)} \begin{bmatrix} \cos^2\theta_a + \cos^2\theta_b & -\tfrac{1}{2}(\sin 2\theta_a + \sin 2\theta_b) \\ -\tfrac{1}{2}(\sin 2\theta_a + \sin 2\theta_b) & \sin^2\theta_a + \sin^2\theta_b \end{bmatrix} \tag{19}$$

Note that the planar GDOP is

$$TR(H^T H)^{-1} = \frac{2}{\sin^2(\theta_a - \theta_b)} \tag{20}$$

This is independent of the choice of coordinate systems and is minimum (best) when $\theta_a - \theta_b = \pm 90^\circ$. When $\theta_a - \theta_b = \pm 90^\circ$, note that the off-diagonal terms in $(H^T H)^{-1}$ are zero, indicating a lack of cross coupling of errors. The geometrical interpretation of these results is clear. The point $\underline{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ is being determined by the intersection of two circles of radii z_a and z_b respectively, centered at $\underline{x}_a$ and $\underline{x}_b$. Small errors in measuring z_a or z_b have the least effect when the circles intersect at right angles. If $\theta_a - \theta_b$ is small so the circles graze each other at a shallow angle, then small errors in z_a or z_b give large errors in $\underline{x}$. An infinite error results if $\theta_a = \theta_b$. This means there is no intersection at all if $z_a \neq z_b$ and no unique intersection if $z_a = z_b$.

Some specific numerical cases related to Figure 1 are informative. Suppose there are two candidate sets of measurements

$$\text{Set 1: } \theta_a = 90^\circ,\ \theta_b = 0 \quad , \ H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\text{Set 2: } \theta_a = 63.435, \theta_b = -45^\circ \ , \ H = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

From Eq. (20) the trace of the GDOP matrix for Set 1 is 2.00. This is the minimum possible value and Set 1 is therefore optimal in terms of GDOP. Set 2 has a corresponding suboptimal value of 2.222. The optimality of Set 1 is expected since $\theta_b - \theta_a = -90^\circ$. For Set 2, $\theta_b - \theta_a = -108.435^\circ$ so the unit measurement vectors are not quite orthogonal, giving the suboptimal GDOP.

Suppose that these two sets of measurements are to be compared in terms of the noise weighted GDOP matrix of Eq. (10). For Set 1, since $H = I$,

$$E(\tilde{\underline{x}}\ \tilde{\underline{x}}^T) = R$$

Suppose that Set 1 has $R = \begin{bmatrix} 4.5 & 0 \\ 0 & 2.5 \end{bmatrix}$ and

Set 2 has $R = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$

Then the trace of Eq. (10) for Set 1 has a value of 7. For Set 2, the corresponding value is 20/3 = 6.667. A very modest decrease in noise variances was sufficient to overcome the GDOP advantage and make Set 2 the best choice in terms of the weighted least-squares criterion.

In order to demonstrate the effect of a priori information, assume that

$$P_o = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$$

and compare measurements Sets 1 and 2 in terms of the trace of the a posteriori covariance in Eq. (13). To nullify effects of different measurement noise levels, assume that

$$R = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$$

for both measurement sets. Straightforward calculation of Eq. (13a) shows that $\mathrm{TR}[\mathbf{M}(\tilde{\underline{x}}\,\tilde{\underline{x}}^T)]$ is 2.823 for Set 1 and 2.697 for Set 2. Thus Set 2 gives the smaller a posteriori error volume even though its GDOP value is inferior to Set 1 and even though the same measurement accuracies are assumed in both cases. These simple results suggest that minimum GDOP does not guarantee minimum a posteriori position uncertainties in many cases. Similar results are found in realistic GPS examples to be presented later.

2.4 Minimum Variance Estimation

If instead of seeking to minimize a selected cost function such as found in Eq. (2), (8) or (11), one seeks the estimate $\hat{\underline{x}}$ which minimizes the error variance $E\{(\underline{x}-\hat{\underline{x}})^T(\underline{x}-\hat{\underline{x}})\}$, then the minimum variance estimator will be obtained. Under quite general conditions

the solution is the conditional mean

$$\hat{\underline{x}} = E\{\underline{x}|\underline{z}\} \tag{21}$$

If it is assumed that the state $\underline{x}$ and the measurement $\underline{z}$ are jointly distributed gaussian random variables, or if the best linear estimator is sought regardless of the form of the distribution functions, then Eq. (21) is known to yield

$$\hat{\underline{x}} = E\{\underline{x}\} + P_{xz}P_{zz}^{-1}[\underline{z} - E\{\underline{z}\}] \tag{22}$$

If the linear measurements of Eq. (1) are assumed, it is well known that Eq. (22) reduces to Eq. (12b), with $\underline{x}_o = E\{\underline{x}\}$. The corresponding a posteriori error covariance matrix is given by Eq. (13b), so that the weighted least-squares and minimum variance solutions are the same under appropriate conditions.

It will be useful for future discussions on time-sequential measurements to list the other well-known equations which make up the recursive form of the minimum variance estimator known as the Kalman filter.[5] The notation used earlier, such as in Eq. (13), with subscript zero indicating conditions just prior to a measurement and subscript one indicating conditions just after, is inadequate when time-sequences are considered. Two time indices are useful. The first index indicates the time at which the estimate applies and the other indicates the most recent measurement used in forming the estimate. Equation (12b) becomes, in this standard Kalman filter notation

$$\hat{\underline{x}}(k+1|k+1) = \hat{\underline{x}}(k+1|k) + P(k+1|k)H(k+1)^T[H(k+1)P(k+1|k)H^T(k+1) + R(k+1)]^{-1}[\underline{z}(k+1) - H(k+1)\hat{\underline{x}}(k+1|k)] \tag{23}$$

The covariance update equation (13b) becomes

$$P(k+1\ k+1) = P(k+1|k) - P(k+1|k)H^T(k+1)[H(k+1)P(k+1|k)H^T + R(k+1)]^{-1}H(k+1)P(k+1|k) \tag{24}$$

The two new equations give the extrapolation between measurement times, and are

$$\hat{\underline{x}}(k+1|k) = \Phi(k+1,k)\hat{\underline{x}}(k|k) + \Gamma(k+1,k)E\{\underline{w}(k)\} \tag{25}$$

and $P(k+1|k) = \Phi(k+1,k)P(k|k)\Phi^{T}(k+1,k) + Q(k)$ (26)

where $\Phi(k+1,k)$ is the state transition matrix

$E\{\underline{w}(k)\}$ is the mean plant noise, often zero

$\Gamma(k+1,k)$ is the input coupling matrix

and $Q(k)$ is the covariance matrix of the plant noise $\Gamma(k+1,k)\underline{w}(k)$

3. SCOPE OF THE STUDY

This research study addresses three major tasks, as outlined below.

3.1 Measurement Selection Methodology

Based upon the results of the simple example in Section 2.3, and previous experience, it is known that the traditional GDOP does not always indicate which sets of potential measurements are optimal. In addition to the geometry effects, which are accounted for by GDOP, nonuniform measurement noise and a priori state information is taken into account and a scalar GDOP-like indicator of static measurement effectiveness is presented.

When a sequence of measurements are to be taken over a period of time, the dynamics of the system and the operational scenarios become factors in measurement selection. This aspect of measurement selection methodology is also addressed.

If a portion of the primary navigation system is degraded because of jamming or other reasons, then secondary or alternate sensors must be considered. This aspect of the measurement selection strategy is addressed, using suboptimal filtering and partitioned state vector concepts.

3.2 Analytical Tests

The methods which result from the task of Section 3.1 are evaluated to the extent practical by analytical methods. Simple models are used for gaining insight, and realistic models are also analyzed. Numerical results are also presented for realistic GPS geometries in some cases. It is pointed out the large scale simulations were explicitly excluded from this study because of the costs that would have been involved in developing the necessary capability which is already in existence at the Air Force Avionics Laboratory.

3.3 Specification of Simulation Cases and Conditions

As mentioned in Section 3.2, the intent of this study was not to duplicate existing simulation capabilities. This study does suggest certain types of cases and conditions that ought to be considered, via simulation, in further evaluation of the concepts developed here.

4. GENERALIZATIONS AND EXTENSIONS OF GDOP

There is absolutely no question that the GDOP concept has played, and will continue to play a major role in the design of many navigation systems. It is simple to use and has modest computational requirements. It characterizes the geometrical aspects of a navigation satellite network very nicely. In the case of satellite network design, the basic requirement is to assure that a wide variety of geographically dispersed users, with a wide variety of sensor accuracies and a priori knowledge are not prevented from solving their navigation problem because of weak geometry.

It is when GDOP is proposed as a means for an individual user to select measurements which are best under his particular circumstances that room exists for improvements, generalizations and extensions. A major part of this study was devoted to the question of how to best select measurements at a given point in time. This static measurement selection problem, so called because only one point in time is considered, is discussed in this section.

To begin, a definition of what is meant by "best" is desirable. The point of view is assumed here that best means the smallest overall mean square position error. In the GPS problem we are concerned with three position components and a clock bias, which converts to an equivalent range error. So for the most part a four-component state vector will be of ultimate interest. Making the usual assumptions about mean values being zero, attention turns to the second statistical moments of the estimation errors, i.e., the covariance terms. Subject to the usual assumptions, Eq. (13a) or (13b) or Eq. (24) tells the entire story. Given an a priori state uncertainty P_o, and a set of measurements characterized by a geometry matrix H and a noise covariance matrix R, the a posteriori covariance P_1 is uniquely computable.

It is awkward to use a 4 x 4 matrix to describe the accuracy or effectiveness of a given set of measurements. An attractive alternative is to use a single scalar number to describe the result. In the past, various terms such as circular error probability (CEP) or spherical error probability (SEP) have been used.

The a posteriori covariance matrix P_1 can be viewed geometrically as a hyperellipsoid with principal axes oriented in space according to the directions of the orthonormal eigenvectors ξ_i of P_1 and with the dimensions of the semi-major axes as the eigenvalues λ_i of P_1. The interior of this ellipsoid can be thought of as a one-sigma error volume, and in this respect it is related to CEP and SEP although much more complicated. Two very simple scalar measures derived from P_1 have been used. The determinant of P_1, is equal to the product of the eigenvalues[6]

$$|P_1| = \lambda_1 \lambda_2 \ldots \lambda_n \tag{27}$$

and is therefore related to the volume of the ellipsoid. This measure is not preferred here for two reasons. First, determinants are not as easy to evaluate as the trace, which is used in the next performance measure. Second, and more important, $|P_1| = 0$ does not mean the error is zero. All it means is that one dimension of the hyperellipsoid has degenerated to zero, but very large errors in the other coordinates may still exist. $|P_1|$ is not very descriptive or discriminating in this regard.

The preferred scalar performance measure is based on the trace of P_1

$$TR(P_1) = \sum_{i=1}^{n} P_1^{ii} = \sum_{i=1}^{n} \lambda_i \tag{28}$$

The trace will normally be computed as the sum of the diagonal elements, but it is known that it also equals the sum of the eigenvalues.[6] Therefore $TR(P_1)$ can not equal zero unless <u>all</u> error variance components are zero.

There may be circumstances where it is more appropriate to use a weighted sum of the diagonal terms of P_1 as a performance measure. The vertical component of position error may be more important than the horizontal components in a terrain-following system, for example. By defining a diagonal matrix,

$$W = \begin{bmatrix} \alpha_1 & & 0 \\ & \alpha_2 & \\ & & \ddots \\ 0 & & \alpha_n \end{bmatrix} \tag{29}$$

it is easily verified that

$$TR(WP_1) = \sum_{i=1}^{n} \alpha_i P_1^{ii} \tag{30}$$

For many purposes $TR(WP_1)$ can be dealt with just as easily as $TR(P_1)$, and they are the same when W is the unit matrix.

The trace functions just defined are natural extensions of GDOP. In the remainder of this section several versions of extended GDOP will be considered. Analytical techniques using eigenvalue-eigenvector theory are developed in an attempt to gain insight into optimal measurement selection. Computational considerations prompt the evaluation of some approximate but easy to compute measures. Numerical rankings of realistic measurement possibilities are provided, to show the performance of the various exact and approximate methods.

4.1 Noise-Weighted GDOP

In Section 2.1 it was agreed that the trace of Eq. (7) would be called GDOP. The trace of Eq. (10) is called the noise-weighted GDOP for obvious reasons,

$$\text{Noise-weighted GDOP} = TR(H^T R^{-1} H)^{-1} \tag{31}$$

The intent is to use measurement accuracy information as well as geometry. This performance measure will prevent the selection of very noisy (large R) measurement sets. This rather minor modification to GDOP is often effective. In fact, comparison with Eq. (13a)

shows it to be exactly $TR(P_1)$ for the case where uncertainties in P_o are so large that P_o^{-1} can be neglected. While this circumstance may apply in the early stages of a navigation solution, it is not particularly appropriate for a well-behaved navigation system in steady-state operation. It will be shown in numerical examples later that noise-weighted GDOP can occasionally do worse than traditional GDOP in selection of measurements. More often it is better than GDOP but still suboptimal.

4.2 Incorporation of A Priori Data

Both Eq. (13a) and (13b) give the a posteriori covariance. Eq. (13a) is most similar to the noise-weighted GDOP matrix, differing only by the inclusion of the a priori covariance P_o. However, Eq. (13b) seems most convenient and will be used here.

Consider as the basic indicator of measurement effectiveness

$$TR(WP_1) = TR(WP_o) - TR\{WP_oH^T[HP_oH^T + R]^{-1}HP_o\} \tag{32}$$

Since W and P_o are unaffected by choice of measurement, Eq. (32) can be minimized (with respect to the choice of H and R) by maximizing

$$SCORE = TR\{WP_oH^T[HP_oH^T + R]^{-1} HP_o\} \tag{33}$$

$$= TR\{HP_oWP_oH^T[HP_oH^T + R]^{-1}\} \tag{34}$$

Equation (34) follows from Eq. (33) by using properties of the trace.[6]

The score is the amount of uncertainty reduction achieved from a given measurement set. If the score is computed for each possible measurement set, the maximum score will indicate the optimum selection. When four satellites are to be selected out of seven assumed to be available for measurement, 35 combinations must be evaluated. In earlier studies, up to ten satellites were assumed visible, giving 210 combinations of four. These are large numbers of possibilities to compute and rank. Therefore effective algorithms for, or good approximations to, Eq. (34) are essential. These algorithmic aspects will be pursued in Section 4.4. First, in order to gain insight and

aid in developing efficient rules of thumb, an analytical investigation is carried out using eigenvalue- igenvector theory.

4.3 Eigenvalue-Eigenvector Analysis[6]

The n x n real symmetric a priori covariance matrix has a full set of n real orthonormal eigenvectors $\underline{\xi}_i$ and n corresponding real nonnegative eigenvalues. They satisfy the eigenvalue equation

$$P_o \underline{\xi}_i = \lambda_i \underline{\xi}_i \qquad i = 1,2,\ldots,n \tag{35}$$

The computation of eigenvalues and eigenvectors for a large matrix can be a burden, and is best avoided in any real-time scheme such as measurement selection. However they provide answers to interesting theoretical questions and are therefore considered in this Section.

The matrix P_o can be written in the so-called spectral decomposition form

$$P_o = \sum_{i=1}^{n} \lambda_i \underline{\xi}_i \underline{\xi}_i^T \tag{36}$$

4.3.1 Individual Scalar Measurements.

The effect of a single scalar measurement is considered first.

$$z = \underline{h}\,\underline{x} + n \tag{37}$$

This is a simplified version of Eq. (1) but now z and n are scalars and $\underline{h}$ is a 1 x n row vector. When Eq. (37) is used, the simplified version of Eq. (13b) is

$$P_1 = P_o - P_o \underline{h}^T \underline{h} / (\underline{h}\, P_o\, \underline{h}^T + R) \tag{38}$$

The measurement noise n has the scalar covariance R.

The set of eigenvectors $\{\underline{\xi}_i,\ i = 1,\ldots,n\}$ can be used as a basis for an n-dimensional vector space, so the column vector $\underline{h}^T$ can be expanded as

$$\underline{h}^T = \sum_{i=1}^{n} \alpha_i \underline{\xi}_i \tag{39}$$

where the expansion coefficients are given by the scalar product

$$\alpha_i = \underline{h}\,\underline{\xi}_i = \underline{\xi}_i^T \underline{h}^T \tag{40}$$

Premultiplying Eq. (39) by P_o and using Eq. (35) gives

$$P_o \underline{h}^T = \sum_{i=1}^{n} \lambda_i \alpha_i \underline{\xi}_i \tag{41}$$

Multiplication of Eq. (41) by $\underline{h}$ and using Eq. (40) gives the denominator required in Eq. (38).

$$\underline{h} P_o \underline{h}^T + R = \sum_{i=1}^{n} \lambda_i \alpha_i^2 + R \tag{42}$$

The spectral form of Eq. (38) is found to be

$$P_1 = \sum_i \lambda_i \underline{\xi}_i \underline{\xi}_i^T - \frac{(\sum_i \lambda_i \alpha_i \underline{\xi}_i)(\sum_k \lambda_k \alpha_k \underline{\xi}_k^T)}{\sum_j \lambda_j \alpha_j^2 + R} \tag{43}$$

By separating out those terms in the double summation for which i=k, this becomes

$$P_1 = \sum_{i=1}^{n} \left\{ \lambda_i - \frac{\lambda_i^2 \alpha_i^2}{\sum_j \lambda_j \alpha_j^2 + R} \right\} \underline{\xi}_i \underline{\xi}_i^T$$

$$- \sum_i \sum_{k \atop i \neq k} \left\{ \frac{\lambda_i \lambda_k \alpha_i \alpha_k}{\sum_j \lambda_j \alpha_j^2 + R} \right\} \underline{\xi}_i \underline{\xi}_k^T \tag{44}$$

Orthonormality of the eigenvectors gives

$$TR(\underline{\xi}_i \underline{\xi}_k^T) = TR(\underline{\xi}_k^T \underline{\xi}_i) = \begin{cases} 0 & i \neq k \\ 1 & i = k \end{cases}$$

Therefore the trace of Eq. (44) is

$$TR(P_1) = \sum_{i=1}^{n} \left\{ \lambda_i - \frac{\lambda_i^2 \alpha_i^2}{\sum_j \lambda_j \alpha_j^2 + R} \right\} \tag{45}$$

The score defined in Eq. (33) and (34) is found from Eq. (45) for a scalar measurement and with W = I,

$$\text{SCORE} = \sum_{i=1}^{n} \frac{\lambda_i^2 \alpha_i^2}{(\sum_j \lambda_j \alpha_j^2 + R)} \tag{46}$$

When complete freedom of choice is available regarding the direction of the measurement vector, it is found that Eq. (46) is maximized and Eq. (45) minimized if $\underline{h}^T = \underline{\xi}_i$, the eigenvector with the largest eigenvalue.

$$\min_{\underline{h}} TR(P_1) = TR(P_o) - \frac{\lambda_{max}^2}{\lambda_{max} + R} \tag{47}$$

When the noise covariance R_i is allowed to vary, but the directions of $\underline{h}_i^T$ are restricted to eigenvectors, the optimal measurement is still the direction of maximum eigenvector $\underline{\xi}_i$ if

$$R_i \leq \frac{(\lambda_i - \lambda_j) + R_j(\lambda_i/\lambda_j)}{(\lambda_j/\lambda_i)} \tag{48}$$

for each other candidate λ_j and R_j. Clearly, if $\lambda_i = \lambda_j$ this indicates that the smallest R_i is optimal. In general a less accurate measurement (larger R_i) may be optimal if there is more to be gained ($\lambda_i > \lambda_j$).

Whenever a scalar measurement vector $\underline{h}^T$ is in the direction of any eigenvector, then there is only one nonzero α_i expansion coefficient in Eq. (39). Assume that $\underline{h}^T$ is a unit vector. Then $\alpha_r = 1$, $\alpha_j = 0$ $j \neq r$ and Eq. (44) becomes

$$P_1 = \sum_{i=1}^{n} \left\{ \lambda_i - \left[\frac{\lambda_1^2 \alpha_i^2}{\lambda_r + R} \right] \delta_{ir} \right\} \underline{\xi}_i \underline{\xi}_i^T$$

$$= \sum_{i=1}^{n} \left\{ \frac{\lambda_i R}{(R + \lambda_i \delta_{ir})} \right\} \underline{\xi}_i \underline{\xi}_i^T \tag{49}$$

where $\delta_{ir} = \begin{cases} 1 & i=r \\ 0 & i \neq r \end{cases}$.

The implications of this result are that P_1 has the same eigenvectors as P_o and all the eigenvalues are unchanged except λ_r. In terms of the error hyperellipsoid, the orientation of its axes remains fixed in space. All coordinate dimensions are also the same as before the measurement, except the one in the direction of the measurement is squeezed down from λ_r to $\lambda_r R/(\lambda_r + R)$. If $R = 0$, this dimension flattens to zero. If R is large, virtually no reduction takes place.

A more likely decision to be faced is the selection of the best measurement from among several which are not necessarily in the direction of eigenvectors of P_o. A straightforward numerical decision based on Eq. (34) will be presented in Sections 4.4 and 4.5. Here, a two-dimensional case is considered analytically.

Assume that the eigenvectors and eigenvalues define the error ellipse shown in Figure 2.

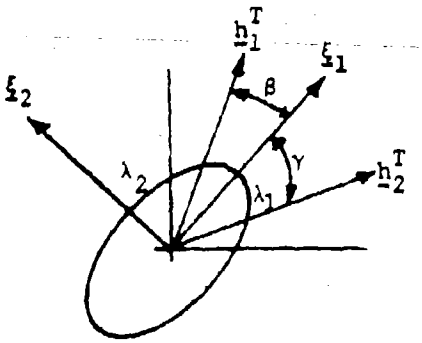

Figure 2. Error Ellipse and Two Candidate Measurement Directions

The candidate measurement vectors make angles β and γ respectively with the semimajor axis. Noise covariances R_1 and R_2 are assumed. The scalar product of two unit vectors is the cosine of their included angle,

$$\underline{h}_1 \underline{\xi}_1 = \cos \beta, \quad \underline{h}_2 \underline{\xi}_1 = \cos \gamma$$

Therefore Eq. (46) can be expressed as

$$S_1 = \frac{\lambda_1^2 \cos^2\beta + \lambda_2^2 \sin^2\beta}{\lambda_1 \cos^2\beta + \lambda_2 \sin^2\beta + R_1}$$

and

$$S_2 = \frac{\lambda_1^2 \cos^2\gamma + \lambda_2^2 \sin^2\gamma}{\lambda_1 \cos^2\gamma + \lambda_2 \sin^2\gamma + R_2} \qquad (50)$$

for the two potential measurements.

In order to determine whether S_1 or S_2 is larger, the difference is couputed, a positive common denominator is omitted and the results manipulated into

$$d = \frac{R_2 - R_1}{\lambda_1} + \left\{\frac{\lambda_2}{\lambda_1}\left(1 - \frac{\lambda_2}{\lambda_1}\right) + \frac{R_1}{\lambda_1}\left[1 - \left(\frac{\lambda_2}{\lambda_1}\right)^2\right]\right\}\sin^2\gamma$$

$$- \left\{\frac{\lambda_2}{\lambda_1}\left(1 - \frac{\lambda_2}{\lambda_1}\right) + \frac{R_2}{\lambda_1}\left[1 - \left(\frac{\lambda_2}{\lambda_1}\right)^2\right]\right\}\sin^2\beta \qquad (51)$$

A positive d indicates that choice 1 is best. Several special cases are of interest. If $\lambda_2/\lambda_1 = 1$ so the ellipse is actually a circle, then the measurement with the smallest R_i is best regardless of the angles γ and β. If λ_2/λ_1 is negligible, so we have an extremely eccentric ellipse, it can be shown that choice 1 is best if

$$R_1 \cos^2\gamma < R_2 \cos^2\beta \qquad (51)$$

One other special case illustrates the trade-offs between eccentricity, off-axis angles and noise levels. Let $\gamma = 0$ so that $\underline{h}_2^T$ is along the maximum eigenvector. It can be shown that d remains positive, so that the off-axis measurement $\underline{h}_1^T$ is still best if

$$|\sin\ \beta| < \sqrt{\frac{(R_2 - R_1)/R_2}{(\lambda_2/R_2)(1 - \lambda_2/\lambda_1) + \left[1 - (\lambda_2/\lambda_1)^2\right]}} \qquad (52)$$

If $R_2 = R_1$, then any nonzero β is suboptimal. If $R_2 > R_1$ so that measurement 1 is more accurate, then with $\lambda_2/\lambda_1 = 1$ Eq. (52) indicates measurement 1 is optimal for any β. For a nontrivial example, let $R_2 = 2R_1$, $\lambda_1 = 3\lambda_2$ and $\lambda_2/R_2 = 1$. Then Eq. (52) indicates that β

can be as much as 34.5° before the effectiveness of measurement 1 becomes less than measurement 2. Notice that for these same parameter values, Eq. (51) overestimates the β limit as 45°. Even so, the noise-level/cosine-squared relationship is a fair rule-of-thumb for eccentric error ellipses. It should predict the correct trends in higher dimensional error hyperellipsoids as well. When the eccentricities are small (nearly equal uncertainty in all directions, then the best choice is the most accurate measurement.

4.3.2 Vector Measurements

The vector measurement case of Eq. (1) is now considered. In this case the noise covariance R is a matrix rather than a scalar. However, if R is diagonal it is known that the a posteriori covariance matrix P_1 of Eq. (13) can be computed by cycling through Eq. (38) once for each measurement. The rows $\underline{h}_i$ of H are used, one at a time along with the i^{th} diagonal term R_i of R. When R is not diagonal, an orthogonal transformation on Eq. (1) can be used to give

$$A\underline{z} = AH\underline{x} + A\underline{n} \tag{53}$$

or

$$z' = H'\underline{x} + \underline{n}' \tag{54}$$

The modified measurement noise $\underline{n}'$ has a covariance matrix

$$E\{\underline{n}'\underline{n}'^{T}\} = E\{A\underline{n}\,\underline{n}^{T}A^{T}\}$$

$$= ARA^{T} = R' \tag{55}$$

It is always possible to select the matrix A so that R' is diagonal. Then Eq. (54) applies in place of Eq. (1), and the noise covariance is again diagonal. For this reason the assumption of a diagonal R matrix is made throughout the rest of the eigenvalue-eigenvector analysis.

Let the eigenvalues of P_o be numbered such that $\lambda_1 > \lambda_2 > \lambda_3 > \ldots > \lambda_n$. Assume that R_1 satisfies Eq. (48) for all j, so that $\underline{h}_1$ is the best single measurement. After processing $\underline{h}_1$, the eigenvalues of the new covariance matrix will be

$$\frac{\lambda_1 R_1}{\lambda_1 + R_1'}, \lambda_2, \lambda_3, \ldots, \lambda_n \tag{56}$$

Of this set, λ_2 will now be the largest if

$$R_1 < \frac{\lambda_1 \lambda_2}{\lambda_1 - \lambda_2} \tag{57}$$

In fact, if all R_i are equal so that Eq. (48) is automatically satisfied whenever $\lambda_i > \lambda_j$, then the optimal measurement set would consist of $\underline{h}_i^T = \underline{\xi}_i$ provided that each R_i satisfies

$$R_i < \frac{\lambda_i \lambda_n}{\lambda_i - \lambda_n} \tag{58}$$

Eq. (58) would assure that each eigenvalue in turn is reduced to a value less than the original smallest one, λ_n. If these conditions were all met, the final P_1 would satisfy

$$TR(P_1) = \sum_{i=1}^{n} \frac{\lambda_1 R_i}{\lambda_i + R_i} \tag{59}$$

Eq. (59) gives the smallest possible $TR(P_1)$ for the case where all R_i are equal, and thus provides a limiting bound on measurement performance.

In actual fact, conditions (57) or (58) are often not satisfied, meaning that a smaller final error could be achieved by using a given measurement direction more than once.

As a simple example, let

$$P_o = \begin{bmatrix} 10 & 0 \\ 0 & 1 \end{bmatrix}$$

and assume that three measurement possibilities exist:

$$\underline{h}_1 = [1 \quad 0] \quad , R_1 = 2$$

$$\underline{h}_2 = [0 \quad 1] \quad , R_2 = 1$$

$$\underline{h}_3 = [-1 \quad 0] \quad , R_3 = 2$$

Both $\underline{h}_1^T$ and $\underline{h}_3^T$ are in the direction of eigenvector $\underline{\xi}_1 (\lambda_1 = 10)$, while $\underline{h}_2^T$ is in the direction of $\underline{\xi}_2 (\lambda_2 = 1)$. Eq. (48) indicates that $\underline{h}_1$ (or $\underline{h}_3$) is the best single measurement. If it is used, the first eigenvalue decreases to 5/3 (See Expression (56)). This is

not less than λ_2, and Eq. (57) is not satisfied. The best second measurement is to use $\underline{h}_1$ again (or if questions about independent measurement noise are of concern, use $\underline{h}_3$ which is the opposite direction.) Skipping some details, the final P_1 matrix for two sets of measurements are

$$\underline{h}_1, \underline{h}_3 : P_1 = \begin{bmatrix} \frac{10}{11} & 0 \\ 0 & 1 \end{bmatrix} \quad TR(P_1) = \frac{21}{11} = 1.9091$$

GDOP = ∞

Noise-weighted GDOP = ∞

$$\underline{h}_1, \underline{h}_2 ; P_1 = \begin{bmatrix} \frac{10}{6} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \quad TR(P_1) = 13/6 = 2.1667$$

GDOP = 2

Noise-weighted GDOP = 4

This example shows that a 13% smaller error is achieved by picking the poorest GDOP and the noisiest measurements. The noise-weighted GDOP is also infinite for the optimal measurement set. This example was constructed to make the point about the inability of GDOP and noise-weighted GDOP to select the best measurements. But it is not that an unusual an example, and it is easy to construct many others of similar behavior. The same phenomenon occurred by itself when numerically processing GPS-type four dimensional data. That is, the worst available GDOP can give the best performance.

It is pointed out that in realistic cases with a fixed finite number of $\underline{h}_1^T$ vectors to choose from, it is not generally true that the best pair of measurements, contains the single best measurement as one of its members. Likewise the best set of four measurements need not include the best pair as a subset. The results in preceding paragraphs are special cases where all measurements are in the direction of the eigenvectors of P_o.

While these are optimal results they may not be among the available options in a real case.

One last topic is illustrated nicely using eigenvalue-eigenvector concepts--the notion of diminishing returns. In the preceding example one measurement in the ξ_1 direction reduced λ_1 from 10 to 5/3. A second measurement in the same direction, and with the same $R_1 = 2$ further reduced λ_1 to 10/11.

Repeated application of the same scalar measurement generates a decreasing sequence of eigenvalues (dimension of the error ellipse) at the k^{th} iteration

$$\lambda_i^{(k+1)} = \frac{\lambda_i^{(k)} R_i}{\lambda_i^{(k)} + R_i} \tag{60}$$

These satisfy

$$0 \leq \lambda_i^{(k+1)} \leq \lambda_i^{(k)} \leq \lambda_i^{(o)} \tag{61}$$

If these inequalities are normalized by the initial value $\lambda_i^{(o)}$, one obtains

$$0 \leq \frac{\lambda_i^{(k+1)}}{\lambda_i(k)} \leq 1 \tag{62}$$

A perfect measurement gives $\lambda_i^{(k+1)}/\lambda_i^{(k)} = 0$ while a useless one gives this ratio a value of one. If initially $R_i = \lambda_i^{(o)}$, then

$$\lambda_i^{(1)}/\lambda_i^{(o)} = \frac{1}{2} \quad , \quad \lambda_i^{(2)}/\lambda_i^{(1)} = \frac{2}{3} \quad ,$$

$$\lambda_i^{(3)}/\lambda_i^{(2)} = \frac{3}{4} \quad , \ldots , \quad \lambda_i^{(k)}/\lambda_i^{(k-1)} = \frac{k}{k+1}$$

This sequence rapidly approaches unity, indicating that further measurements in that direction are of little or no value. The error ellipsoid has been flattened to almost zero in that coordinate direction. Actually, this demonstration of diminishing returns applies exactly to scalar measurments in the direction of an eigenvector. However, a similar effect is actually observed in higher dimensions as well. All other things being equal, some sort of alternation of measurements seems better than a repetition of the same set.

Using the error ellipsoid concept for heuristic justification, the best strategy seems to be to squeeze down on the largest dimension of the ellipsoid, then another large dimension, and so on, always trying to deform the ellipsoid into the smallest possible sphere. Time-squences of measurements are discussed further in Section 5.

4.4 Algorithmic Approach to Measurement Selection

The basis for this section is Eq. (34). Consider the typical GPS application where measurements to four satellites constitute one measurement set. Assume that there are N satellites visible from which the best four are to be selected. (N would typically be 7 or less). Define the N x 4 matrix H_c whose rows are the vectors associated with a measurement to each satellite. If the i^{th} satellite has azimuth angle A and elevation angle E, the i^{th} row is

$$H_i = [\cos E \cos A \quad \cos E \sin A \quad \sin E \quad 1] \tag{63}$$

Also define the N x N matrix R_c as the covariance matrix for all N noise components jointly. H_c and R_c would have to be given data, along with P_o and the weighting matrix W. Two more symmetric N x N matrices are computed from the given data. They are

$$\begin{aligned} B &= H_c \, P_o \, H_c^T + R_c \\ C &= H_c \, P_o \, W \, P_o \, H_c^T \end{aligned} \tag{64}$$

The score associated with any number of measurements, from 1 through N, can be obtained by computations on Eq. (64). For example, if all N measurements were to be used (and this would give the smallest possible a posteriori error) then Eq. (34) gives

$$\text{Score} = TR(CB^{-1}) \tag{65}$$

but then there are no options to rank. In general, fewer than N measurements will be processed. To indicate the number of measurements, as well as which ones, a subscripted S will be used for the score. For example S_3 means one measurement to satellite 3 is made, S_{ij} means a pair of measures, to satellites i and j are used, and

S_{ijkl} means four measurements are used, to satellites i,j,k and l. These various scores are computed as in Eq. (65), but subsets of the matrices C and B are used.

4.4.1 Single Measurement Scores

To determine the relative ranking of N possible single measurements, the following N scores must be computed and ranked.

$$S_i = C_{ii}/B_{ii} \qquad i = 1,N \tag{66}$$

where C_{ii} and B_{ii} are the i,i components of C and B of Eq. (64).

4.4.2 Pairs of Measurements

In the case of measurements to satellite i and satellite j, the 2 x 2 submatrices formed from the intersections of rows i and j with columns i and j of B and C are formed and used in Eq. (65) instead of the full N x N matrices. The explicit form of the 2 x 2 inverse is easily computed, giving the score

$$S_{ij} = [B_{ii}C_{jj} + B_{jj}C_{ii} - 2B_{ij}C_{ij}]/[B_{ii}B_{jj} - B_{ij}^2] \tag{67}$$

There are $\frac{N!}{2(N-2)!} = \frac{N(N-1)}{2}$ sets of i,j pairs that would need to be computed and ranked. The optimum i,j pair corresponds to the largest value of S_{ij}. It is convenient to store the results of Eq. (66) along the diagonal of an N x N triangular matrix F, and the results of Eq. (67) in the i,j^{th} off diagonal location. Note from Eq. (67) that $S_{ij} = S_{ji}$. This triangular matrix F will be used in an approximate ranking scheme for four measurements, in Section 4.2.4.

4.4.3 The General Case, Exact Formulation

Although these methods apply for any number of measurements from 1 to N, GPS is most interested in 4 satellite fixes. Attention is therefore focused on evaluation of S_{ijkl}. For each set of ijkl values a pair of 4 x 4 matrices $\hat{B}$ and $\hat{C}$ are formed from B and C as the elements at the intersections of rows and columns i,j,k,l. No simple explicit formulas are available for the 4 x 4 inverses that are required by

$$S_{ijkl} = TR(\hat{C}\,\hat{B}^{-1}) = TR(\hat{B}^{-1}\,\hat{C}) \tag{68}$$

It is inefficient to actually compute the 4 x 4 inverse and then multiply by $\hat{C}$ and then take the trace. It is more efficient to define an unknown matrix D by the set of simultaneous equations

$$\hat{B}\,D = \hat{C} \tag{69}$$

and then use a linear equation solving routine such as gaussian elimination to solve for D. Actually only the diagonal elements of D need be calculated since

$$S_{ijkl} = TR(D) = \sum_{\alpha=1}^{4} D_{\alpha\alpha} \tag{70}$$

This is the approach used in the numerical studies of Section 4.2.5. An alternate method of computation is available, using Bôcher's formula[7].

$$TR(\hat{C}\,\hat{B}^{-1}) = -\frac{a_3}{a_4}\,TR(\hat{C}) - \frac{a_2}{a_4}\,TR(\hat{C}\,\hat{B})$$

$$-\frac{a_1}{a_4}\,TR(\hat{C}\,\hat{B}^2) - \frac{1}{a_4}\,TR(\hat{C}\,\hat{B}^3) \tag{71}$$

where

$$\begin{aligned}
a_1 &= -TR(\hat{B}) \\
a_2 &= -\tfrac{1}{2}\left[a_1 TR(\hat{B}) + TR(\hat{B}^2)\right] \\
a_3 &= -\tfrac{1}{3}\left[a_2 TR(\hat{B}) + a_1 TR(\hat{B}^2) + TR(\hat{B}^3)\right] \\
a_4 &= -\tfrac{1}{4}\left[a_3 TR(\hat{B}) + a_2 TR(\hat{B}^2) + a_1 TR(\hat{B}^3) + TR(\hat{B}^4)\right]
\end{aligned} \tag{72}$$

While this algorithm appears simple, the necessity of computing the various matrix powers and products may actually make it less efficient than Eq. (70). No detailed comparison of these methods was carried out during this study.

4.4.4 Approximation Methods

Because of the computational burden involved in evaluating the exact score for a large number of candidate measurement

combinations, the desire to find simple approximations is quite strong. Traditional GDOP and noise-weighted GDOP can rightfully be considered as examples of approximations to what is really wanted. Various other approximations have been investigated during this study. One series of approximations grows out of the inequalities[8]

$$TR(\hat{C})\lambda_{min}(\hat{B}^{-1}) \leq TR(\hat{C}\,\hat{B}^{-1}) \leq TR(\hat{C})\lambda_{max}(\hat{B}^{-1}) \tag{73}$$

or

$$\frac{TR(\hat{C})}{\lambda_{max}(\hat{B})} \leq TR(\hat{C}\,\hat{B}^{-1}) \leq \frac{TR(\hat{C})}{\lambda_{min}(\hat{B})} \tag{74}$$

These inequalities provide bounds on the true score, and depend on finding the maximum and minimum eigenvalues of $\hat{B}$ rather than requiring that $\hat{B}^{-1}$ be computed. The bounds provided by Eq. (73) and (74) are often so loose as to be of little value in properly ranking measurements. In addition, the determination of eigenvalues is in itself a computational nuisance. In an attempt to eliminate the eigenvalue dependency, another inequality was considered,

$$n\,\lambda_{min} \leq TR(\hat{B}) \leq n\,\lambda_{max} \tag{75}$$

Eq. (75) combined with Eq. (74) suggests that perhaps

$$TR(\hat{C})/TR(\hat{B})$$

might have merit as an extremely simple to compute method of ranking measurements. Numerical experimentation along these lines generally proved disappointing.

It was known from the outset that the optimum set of four measurements, say to satellites i,j,k, and l, need not contain the best pair of measurements, which might involve satellites i and m for example. However, it was felt that all six pairs that can be made from four indices i,j,k,l ought to give pretty good pair scores. That is, if S_{ijkl} is the largest of all possible 4-measurement scores, then S_{ij}, S_{ik}, S_{il}, S_{jk}, S_{jl} and S_{kl} should have reasonably high scores when compared with all pair scores. Because of the ease with which all S_{ij} and S_i scores can be exactly computed and stored

in the triangular matrix F of Section 4.2.2, it would be attractive to approximate S_{ijkl} in terms of elements of F. Considerable time was spent in trying to discover a good approximation of this type. Theoretical, empirical and ad hoc approaches to the problem were tried. The best result, whose derivation can be found in[9] is

$$\hat{S}^2_{ijkl} = S^2_{ij} + S^2_{kl} + S^2_{ik} + S^2_{jk} + S^2_{il} + S^2_{jl} - 2[S^2_i + S^2_j + S^2_k + S^2_l] \quad (76)$$

where $\hat{S}_{ijkl}$ indicates an "approximation to" S_{ijkl}. This approximation will be tested, along with the exact result and GDOP results, in the next Section.

4.5 Numerical Examples of Measurement Selection

This Section evaluates GDOP, noise-weighted GDOP, the exact score (with or without the weighting matrix W) and the approximate score $\hat{S}$. The evaluation consists of seeing how well each of these methods is able to rank 35 potential measurement combinations of 4 out of seven satellites with realistic GPS geometries.

The first task is selection of input data. The geometry (H matrices) were abundantly available.[10,11] However, the measurement noise covariance matrices R were not. The more usual way of specifying measurement noise is in terms of jamming power to signal power (J/S) ratios, in decibel units.

The problem of converting J/S values to noise covariance values for the R matrix is nontrivial and depends upon the type of jamming, the receiver characteristics (such as bandwidth, etc.). By assuming wide-band jamming noise, (essentially white noise) and narrow-band signals, along with other drastic over-simplifications, it appears that the variance in a given signal is related to J/S according to

$$R_{ii} = \alpha[1 + \beta\, 10^{(J/S)+\ 10}] \quad (77)$$

where α is the nonjamming variance and β is a function of the receiver bandwidth. Many J/S values seem to be between about 25 and 90 db.

Values of $\alpha = 2.5$ and $\beta = .01$ were used, as well as $\alpha = 25$, $\beta = .005$. The scale factor α is not needed, if only rankings of $(H^T R^{-1} H)^{-1}$ are being made. However, the more general Eq. (13) or (34) require the scale factor be known relative to values from P_o. The off-diagonal entries of the R matrix are related to the angular separation between measurement directions, and the antenna beam width. The correlation between two measurements was programmed as

$$\rho_{ij} = \frac{1}{2} \cos^2 \theta_{ij} \quad \text{if} \quad \theta_{ij} \leq 40^\circ$$

$$= 0 \qquad \text{otherwise} \tag{78}$$

A few cases of nonpositive definite R matrices have cropped up using this model. It is not implied that this noise model is optimum or even totally correct. It is felt that exact numbers are not crucial to the main points of this study. When better noise models become available, they can be used with the selection methods developed here.

The other key input data is P_o, a 4 x 4 matrix. In order to proceed with the study a limited parametric approach was taken regarding the size of the eigenvalues λ_i of P_o. Values were grouped into the classifications large, typical and small. The eccentricity of the P_o ellipsoid was varied also, in terms of ratios of λ_i/λ_j. The values ranged from highly eccentric to spherical. A complete parametric study was not carried out. Six different P_o matrices were tested, along with 8 different sets of satellite geometries, several sets of J/S values in some cases, and a few tests with a non-unit W matrix. About one hundred combinations were tested. Eleven complete sets of results, each consisting of two pages of computer printout, were presented and discussed in[12]. Some of these results are also discussed in[9]. Table 1 summarizes the key features of the numerical experiments. Case numbers shown agree with earlier reports. In Table 1 r_T indicates the true rank, based on $TR(WP_1)$ of the measurement combination chosen as best by the various methods. The value r_o is the rank, as determined by the various methods, of

Case No.	Brief Descrip.	GDOP	N.W. GDOP	$\hat{S}$	S or $TR(P_1)$	S or $TR(WP_1)$	
5	Typical P_o	29,16	2,11	3,5	1,1	1,1	r_T, r_o
	variable J/S	545.8	412.6	413.9	410.4	410.4	e
	W=I	33	.6	.95	0	0	%
11	Large, more	3,33	6,33	4,2	1,1	1,1	r_T, r_o
	eccentric P_o	1719	1747	1720	1704	1704	e
	variable J/S	.9	2.5	.9	0	0	%
	W=I						
20	Small P_o	5,35	2,13	1,1	1,1	1,1	r_T, r_o
	variable J/S	41.4	40.5	39.2	39.2	39.2	e
	W=I	5.6	3.3	0	0	0	%
24	Large P_o	1,1	2,2	13,21	1,1	1,1	r_T, r_o
	equal J/S	361.8	362.9	677.5	361.8	361.8	e
	W=I	0	.3	87	0	0	%
33	Same as case	28,15	2,8	3,6	1,1	1,1	r_T, r_o
	5 except	455	280	282	279	279	e
	increased	63	.5	1	0	0	%
	correlations						
	in P_o						
44	Typical	16,11	5,3	3,2	1,1	1,1	r_T, r_o
	sized P_c,	452	387	368	359	359	e
	but spheric.	20	7.9	2.6	0	0	%
	variable J/S						
	W=I						
51	Small, highly	22,17	7,8	5,5	1,1	1,1	r_T, r_o
	eccentric P_o	281	202	200	195.7	195.7	e
	variable J/S	43.7	3.2	2.2	0	0	%
	W=I						

Table 1; Summary of STATIC Measurement Selection Experiments (continued)

Case No.	Brief Descrip.	GDOP	N.W. GDOP	$\hat{S}$	S or $TR(P_1)$	S or $TR(WP_1)$	
60	Same as	33,16	4,11	2,7	1,1	1,1	r_T, r_o
	case 5 but	1022	708	707	698	698	e
	W_{ii}=2.5, 1., .25, .25	46	1.5	1.3	0	0	%
61	Same as	21,9	1,1	6,3	4,2	1,1	r_T, r_o
	case 5 but	357	320.4	327.7	326.6	320.4	e
	W_{ii}=.25, 2.5, .25, 1	11.6	0	2.3	1.9	0	%
62	Same as	19,16	2,11	3,5	1,1	1,1	r_T, r_o
	case 5 but	309	259	259.2	258.9	258.9	e
	W_{ii}=.25, .25, 2.5, 1	19.3	0	.1	0	0	%
63	Same as	33,16	4,11	2,5	1,1	1,1	r_T, r_o
	case 5 but	461	333.8	332.8	329.3	325.3	e
	W_{ii}=1, .25, .25, 2.5	40	1.4	1	0	0	%
		35.9	13.5	9.6	2.4	2	Avg. r_T+r_o

Table 1: Summary of STATIC Measurement Selection Experiments (concluded)

the true optimum choice. For example, in Case 5 GDOP selected as best a combination which was actually only 29^{th} best out of 35. GDOP placed the true optimum in 16^{th} place. The values of r_T and r_o are all one in the $TR(WP_1)$ column, because the top choice of minimum $TR(WP_1)$ or maximum S is by definition the over-all optimum. The values e in Table 1 are the results of $TR(WP_1)$ when the top choice measurement combination is used. In Case 5, GDOP selected a measurement set which was really only 29^{th} best, and this gave a value of $TR(WP_1) = 545.8$, compared with the optimal choice which gave 410.4. Thus using GDOP caused a 33% larger $TR(WP_1)$ than the optimal. All percentages above optimal are given in Table 1.

Note that Case 20 is an example where the true optimal choice has the worst of the 35 GDOP values for this geometry. Since the a priori errors are very small, as evidence by the small P_o, there is little to be gained from the measurements. The absolute differences of the results are therefore small no matter which selection method is used.

Case 24 has a large P_o and all J/S values are equal. This is the situation where GDOP is expected to do well, and it does. The noise-weighted GDOP does slightly worse even though all diagonal terms in R are equal. This is due to measurement correlations, which cause nonzero off diagonal terms. Using R^{-1} as a weighting term hurt more than it helped in this case, but the differences are small. Case 24 also illustrates the occasional example where $\hat{S}$ is not a good discriminator of measurement effectiveness. Cases 60 through 64 use the same input data as Case 5, except various W weights are used. The trace of W is held at 4 in all cases.

Some generalizations can be stated about the above results. The use of GDOP to select measurement combinations involves computing the inverse of a 4 x 4 matrix (or the equivalent computation) for each potential measurement set. This procedure generally does not give the best selection. It becomes better as P_o becomes large and when the measurement noise levels are nearly equal.

Adding the inverse of the measurement noise covariance to GDOP gives what is called here noise-weighted GDOP. It usually gives better results than GDOP, if P_o is large and if the measurement

noise levels are unequal. There are examples where this generalization gives a worse selection than GDOP (Cases 11 and 20).

A reasonably simple approximate ranking scheme uses exact analysis of measurement pairs to predict effectiveness of sets of four measurements. This is the $\hat{S}$ approximation method. It often works quite well, almost always better than GDOP. The occasional failures such as Case 24, were fairly dramatic, and are sufficient reason to consider it unreliable.

The exact analysis of the weighted trace of the a posteriori state covariance is the true optimal method. Its computational burden is not really much higher than the 4 x 4 matrix inversion required by the GDOP computation. Several methods of carrying out this exact computation were discussed in Section 4.2.3.

The preceding analysis and results concludes the consideration of the static measurement selection problem. It is possible to construct scenarios where the over-all best time sequence of measurements is not the one which results from selecting the best set at each possible decision time. For example, if the time needed to acquire and settle on a new combination of satellites is sufficiently long some measurement cycles may be missed. The measurement omissions may contribute more to the time average error than would use of a slightly suboptimal measurement set, without omissions. Some of these issues are considered in the next Section.

5. SYSTEM DYNAMICS AND TIME-SEQUENTIAL MEASUREMENTS

Several versions of the optimal measurement selection problem have been considered in the literature.[8,13,14,15,16] Although this previous work may have indirectly influenced this section, it does not appear that a ready solution to the GPS measurement selection problem is yet available.

This section assumes that a generalization to GDOP is available which indicates at any given time which combination of measurements will yield the smallest mean square position error. The question addressed here is: What is the optimal strategy for selecting a sequence of measurements over some period of time? In particular, does the optimal sequence consist of selecting the instantaneous best, in the generalized GDOP sense, at each decision time? This is a very complicated question whose answer depends upon many parameters, conditions and ground rules as yet unspecified. To begin the investigation, a very simple model is first considered.

5.1 Second-Order Model

Let the system state vector $\underline{x}$ have just two components, position and velocity. Assume that at each measurement time only two choices are available:

$$z_1 = [1 \quad 0]\,\underline{x} + n_1$$

or (79)

$$z_2 = [0 \quad 1]\,\underline{x} + n_2$$

Since z_1 is a direct measurement of position, it will be more effective in reducing position uncertainties.

Likewise z_2 ought to be more effective in reducing velocity uncertainties. It will become clear soon that both z_1 and z_2 affect both position and velocity somewhat, unless the correlation between position and velocity is zero.

Let the state covariance matrix before a measurement be

$$P_o = \begin{bmatrix} \sigma_P^2 & \sigma_{PV} \\ \sigma_{PV} & \sigma_V^2 \end{bmatrix} \tag{80}$$

Then the covariance $P(k|k)$ just after a measurement is computed from Eq. (13) or (24). Let the selection of measurement z_1 be option A. Then, $H = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $R = R_1$

$$\overset{A}{P}(k|k) = \begin{bmatrix} \sigma_P^2 & \sigma_{PV} \\ \sigma_{PV} & \sigma_V^2\left[1 + \frac{\sigma_P^2}{R_1}(1-\rho^2)\right] \end{bmatrix} \Big/ (1 + \sigma_P^2/R_1) \tag{81}$$

Let the choice of z_2 be option B. Then $H = \begin{bmatrix} 0 & 1 \end{bmatrix}$ and $R = R_2$, yielding

$$\overset{B}{P}(k|k) = \begin{bmatrix} \sigma_P^2\left[1 + \frac{\sigma_V^2}{R_2}(1-\rho^2)\right] & \sigma_{PV} \\ \sigma_{PV} & \sigma_V^2 \end{bmatrix} \Big/ (1 + \sigma_V^2/R_2) \tag{82}$$

If the "best" choice is defined as the one which gives the smallest position uncertainty after just one measurement, then a comparison of the 1,1 elements of $\overset{A}{P}(k|k)$ and $\overset{B}{P}(k|k)$ is appropriate. The choice depends on several parameters, even in this simple case. But as $\rho \to 0$ option A clearly is the best choice. This direct measurement of position would be expected to be the best way to reduce position uncertainties in general, (although an unusual combination of ρ, σ_V^2/R_2 and σ_P^2/R_1 could cause this not to be true).

When the "best" choice is defined to mean the one which gives the smallest average position error over the interval T between two measurement times, the decision is more interesting. To investigate this question, the transition matrix Φ is needed, as well as the plant noise covariance matrix Q. A reasonable approximation for Φ in many cases is

$$\Phi = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$$

Then the extrapolated covariance matrix (after making a measurement and then extrapolating up to just prior to the next measurement) is $P(k+1|k)$ computed as in Eq. (26). If option A is selected, the 1,1 component is

$$\overset{A}{P}(k+1|k)_{1,1} = \frac{\{\sigma_P^2 + 2T\,\sigma_{PV} + \sigma_V^2\left[1 + \sigma_P^2(1-\rho^2)/R_1\right]T^2\}}{1 + \sigma_P^2/R_1} \tag{83}$$

If option B is selected, the 1,1 component is

$$\overset{B}{P}(k+1|k)_{1,1} = \frac{\{\sigma_P^2\left[1 + \sigma_V^2(1-\rho^2)/R_2\right] + 2T\,\sigma_{PV} + \sigma_V^2\,T^2\}}{1 + \sigma_V^2/R_2} \tag{84}$$

These two expressions could be compared if the desire is to get the smallest extrapolated position error. However, it seems more appropriate to seek the smallest average error over the interval T. If a straight line variation is assumed between the beginning value (from Eq. (81) or (82)) and the ending value (from Eq. (83) or (84)) then the appropriate thing to do is compare the averages

$$\frac{1}{2}\,T\left[\overset{A}{P}(k|k) + \overset{A}{P}(k+1|k)\right] \overset{?}{\lessgtr} \frac{1}{2}\,T\left[\overset{B}{P}(k|k) + \overset{B}{P}(k+1|k)\right] \tag{85}$$

Carrying out the algebra and subtracting side B from side A gives the decision variable D. Note that the subtraction causes the plant noise terms to cancel.

$$\begin{aligned} D = {} & \left[2\left(\frac{\sigma_P}{\sigma_V T}\right)^2 + 2\left(\frac{\sigma_P}{\sigma_V T}\right)\rho + 1\right](g_A - g_B) \\ & + \sigma_P^2(1-\rho^2)\left[\frac{g_A}{R_1} - \frac{2g_B}{R_2 T^2}\right] \end{aligned} \tag{86}$$

where $g_A = 1/(1 + \sigma_P^2/R_1)$ and $g_B = 1/(1 + \sigma_V^2/R_2)$. $D < 0$ means choice A is best, while $D > 0$ means choice B is best, and $D = 0$ means both choices are equally good.

The trade-off can best be described in terms of Figure 3.

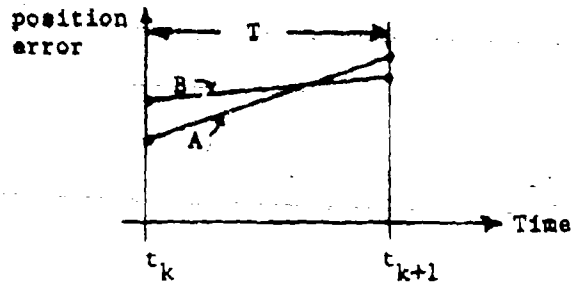

Figure 3. Linear Error Growth

Choice A will normally give a smaller initial position error but a large velocity error. The velocity error determines, at least in part, the slope of the error growth between measurements. Even for the simple problem under consideration, there are six parameters that go into the decision, namely σ_P, σ_V, ρ, T, R_1 and R_2.

From Eq. (86) it is seen that if $\sigma_P^2/R_1 = \sigma_V^2/R_2$ (sort of a signal to noise ratio SNR, then $g_A = g_B$. If in addition $\rho^2 = 1$, or if $R_1 = R_2T^2/2$, then D = 0. Combining the two conditions not involving ρ shows that D = 0 if the SNR are equal and if $\sigma_V^2 T^2 = 2\sigma_P^2$. If $\sigma_V^2 T^2 > 2\sigma_P^2$ then option B is best.

There are many combinations of parameters that lead to choice B being better than choice A. This suggests a partial answer to the question which opened this Section. The choice of measurements which gives the smallest <u>instantaneous position error</u> is not necessarily the one which will give the smallest <u>time-average</u> position error. A very similar system model has been considered[13] and it is shown that if only one type of measurement can be used, the position measurements are optimal. When either position or velocity measurement can be used, it is found that a combination of both is best over a period of time. These results[13] apply for a specific set of parameters equivalent to the six mentioned in the preceding paragraph. An examination of a more realistic model of the GPS system is now presented to show that similar conclusions also apply there.

5.2 More General GPS Model

Let the state vector $\underline{x}$ consist of seven components; three position components plus a time bias form the first four components, and these four have been dealt with extensively in earlier sections. The last three elements of $\underline{x}$ are the user velocity components. The state covariance matrix P is 7 x 7 and will be partitioned as

$$P = \left[\begin{array}{c|c} P_1 & P_2 \\ \hline P_2^T & P_3 \end{array}\right] \tag{87}$$

where P_1 is the 4 x 4 block dealt with in previous sections, P_3 is the 3 x 3 covariance of the velocity terms and P_2 is the 4 x 3 matrix of correlations.

It is assumed that measurements are processed in sets of four. That is, $\underline{z}$ is a four-component vector related to the state by

$$\underline{z} = H\,\underline{x} + \underline{n} \tag{88}$$

The measurement matrix H is 4 x 7 and can also be partitioned as

$$H = \left[H_1 \;\vdots\; 0_{4x3}\right] \tag{89}$$

where H_1 is the 4 x 4 measurement matrix of previous sections. The measurement noise covariance matrix R is also of dimensions 4 x 4.

As the result of processing the $k+1^{th}$ group of four measurements, the covariance is updated by applying Eq. (24), which in partitioned form becomes

$$\left[\begin{array}{c|c} P_1(k+1|k+1) & P_2(k+1|k+1) \\ \hline P_2^T(k+1|k+1) & P_3(k+1|k+1) \end{array}\right] = \left[\begin{array}{c|c} P_1(k+1|k) & P_2(k+1|k) \\ \hline P_2^T(k+1|k) & P_3(k+1|k) \end{array}\right]$$

$$- \left[\begin{array}{c|c} P_1(k+1|k)H_1^T\left[H_1P_1(k+1|k)H_1^T + R\right]^{-1}H_1P_1(k+1|k) & \\ \hline P_2^T(k+1|k)H_1^T\left[H_1P_1(k+1|k)H_1^T + R\right]^{-1}H_1P_1(k+1|k) & \end{array}\right.$$

$$\left.\begin{array}{c|c} & P_1^T(k+1|k)H_1^T\left[H_1P_1(k+1|k)H_1^T + R\right]^{-1}H_1P_2(k+1|k) \\ \hline & P_2^T(k+1|k)H_1^T\left[H_1P_1(k+1|k)H_1^T + R\right]^{-1}H_1P_2(k+1|k) \end{array}\right] \tag{90}$$

It is seen that the upper left hand 4 x 4 partition of Eq. (90) is exactly the same form as was dealt with when considering generalized GDOP performance measures.

In order to analyze time-sequential behavior a state transition matrix $\Phi(k+1,k)$ and a plant noise covariance matrix $Q(k)$ are required. Assume that the motion of the user can be approximated as rectilinear motion over the small time $T = t_{k+1} - t_k$, and that the time bias remains constant except for random effects. Then the 7 x 7 Φ matrix can be approximated by

$$\Phi(k+1,k) = \left[\begin{array}{ccc|c|ccc} & & & 0 & & & \\ & & & 0 & & & \\ & I_3 & & 0 & & I_3\,T & \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hline & & & 0 & & & \\ & 0_{3x3} & & 0 & & I_3 & \\ & & & 0 & & & \end{array}\right] \tag{91}$$

where I_3 and 0_{3x3} are the 3 x 3 unit and null matrices respectively.

If all plant noise effects on position and velocity are attributed to 3 independant components of random acceleration with equal variances σ_a^2, and if the variance of the random time bias drift is $\beta\,\sigma_a^2\,T^2$, then the 7 x 7 Q matrix can be approximated as

$$Q(k) = \left[\begin{array}{ccc|c|ccc} & & & 0 & & & \\ & I_3\,\frac{T^4}{4} & & 0 & & I_3\,\frac{T^3}{3} & \\ & & & 0 & & & \\ \hline 0 & 0 & 0 & \beta T^2 & 0 & 0 & 0 \\ \hline & & & 0 & & & \\ & I_3\,\frac{T^3}{3} & & 0 & & I_3\,T^2 & \\ & & & 0 & & & \end{array}\right] \sigma_a^2 \tag{92}$$

Using these approximations, the covariance propagation equation of Eq. (26) can be written as

$$P(k+1\ k) = \left[\begin{array}{c|c} P_1 + \left[\begin{array}{c} I_3 \\ \hline 0\ 0\ 0 \end{array}\right] P_2^T\,T + P_2 \left[\begin{array}{c|c} I_3 & \begin{array}{c}0\\0\\0\end{array} \end{array}\right] T + \left[\begin{array}{c} I_3 \\ \hline 0\ 0\ 0 \end{array}\right] P_3 \left[\begin{array}{c|c} I_3 & \begin{array}{c}0\\0\\0\end{array} \end{array}\right] T^2 & P_2 + \left[\begin{array}{c} I_3 \\ \hline 0\ 0\ 0 \end{array}\right] P_3\,T \\ \hline P_2^T + P_3 \left[\begin{array}{c|c} I_3 & \begin{array}{c}0\\0\\0\end{array} \end{array}\right] T & P_3 \end{array}\right] + Q(k) \tag{93}$$

The trace of the upper left 4 x 4 partition of P will be investigated, since this is in essence what the GDOP and generalized GDOP approaches use. Define

$$e(k+1|k) = TR\left[P_1(k+1|k)\right] \tag{94}$$

and

$$e(k+1|k+1) = TR\left[P_1(k+1|k+1)\right] \tag{95}$$

Then from Eq. (92) and Eq. (93)

$$e(k+1|k) = e(k|k) + 2aT + bT^2 + \frac{3\sigma_a^2}{4} T^4 + \beta\sigma_a^2 T^2 \tag{96}$$

where

a = TRACE of the first 3 x 3 sub-block of $P_2(k|k)$

$$b = TR\left[P_3(k|k)\right] \tag{97}$$

From Eq. (90)

$$e(k+1|k+1) = e(k+1|k) - S(k+1) \tag{98}$$

where S(k+1) is the "score" associated with the measurement set selected at time k+1, as introduced in Eq. (33) and (34) but with W=I.

To be consistent with Eq. (85) of Section 5.1, potential measurements (defined by choice of H_1 and R) could be compared based on the average

$$\frac{1}{2} T\left[e(k|k) + e(k+1|k)\right] .$$

It can be seen that exactly the same kinds of trade-offs exist here as in Section 5.1. That is, some candidate measurement set may produce the lowest e(k|k) (generalized GDOP) at time k, but may not be optimal because of a larger associated velocity uncertainty as represented by b in Eq. (96) and Eq. (97).

Instead of just looking at a single time step, the average position error over K time-cycles (the planning horizon) ought to be minimized by choice of measurements. That is,

$$\text{minimize } J = \sum_{k=o}^{K} e(k+1|k+1) \tag{99}$$

by considering all admissible measurements, subject to the recursive equation, obtained by comparing Eq. (96) and (98).

$$e(k+1|k+1) = e(k|k) + 2a_k T + (b_k + \beta\sigma_a^2)T^2 + \frac{3\sigma_a^2}{4} T^4 - S(k+1) \qquad (100)$$

The decision variable is the score S(k+1). The formulation of Eq. (99) and (100) looks like a standard type of discrete optimal control problem that would be amenable to solution perhaps using dynamic programming. This is not exactly true, because

(1) the score S(k+1) is in fact a function of $e(k+1|k)$ and thus indirectly $e(k|k)$

(2) the parameters a_k and b_k are not constants but are affected by the choice of measurements

(3) the real crux of the GPS problem is the specification of the set of admissible measurements. Obviously, satellites must be above some minimum elevation angle if they are to be admissible candidates. More than this however, a variety of ground rules could be imposed. For example, once a given satellite is selected, it may be required to stay with it for some minimum time. Also, to allow for acquisition and settling times, a new satellite may not be admissible until after some switch-over delay. Thus, the admissible measurements change with time in some complex way.

It is possible that Eq. (100) might be made a useful approximate model by assuming some typical average values for a_k and b_k (correlations and velocity uncertainties) and by approximating S(k+1) by a measurement effectiveness indicator analogous to the original GDOP which does not depend on the covariance. This type of formulation could then be solved, using dynamic programming for example. This approach would have to be investigated by means of extensive realistic simulations, and is therefore outside the scope of this study.

5.3 A Simple Sequential Example

The same example of Section 2.3 is reconsidered. This could be treated as the 2 x 2 position sub-block of a 4 x 4 system with $\Phi = I$ and $Q = 0$. Alternately it might be repeated measurements on a static system. At each measurement time a choice must be made between measurement sets 1 and 2. Both have the same noise matrix R, and the initial covariance is the same one used in Section 2.3. In that previous section it was shown that set 1 has a slightly better GDOP but set 2 gives a smaller mean square error after one measurement update. Here, all eight sequences of measurements over the first three measurement times are analyzed. Major results are given in Table 2. The initial mean square error in all eight cases is $TR(P_o) = 6$. The successive values of $e(1|1)$, $e(2|2)$ and $e(3|3)$ defined in Eq. (98) or (100) are tabulated.

Measurement Sequence						
1st	2nd	3rd	$e(1\|1)$	$e(2\|2)$	$e(3\|3)$	$\sum_{i=1}^{3} e(i\|i)$
H_1	H_1	H_1	2.823	1.895	1.433	6.151
H_1	H_1	H_2	2.823	1.895	1.352	6.070
H_1	H_2	H_1	2.823	1.907	1.433	6.163
H_1	H_2	H_2	2.823	1.907	1.443	6.173
H_2	H_1	H_1	2.697	1.782	1.352	5.831
H_2	H_1	H_2	2.697	1.782	1.352	5.831
H_2	H_2	H_1	2.697	1.856	1.352	5.905
H_2	H_2	H_2	2.697	1.856	1.433	5.986

Table 2: A Sequence of Three Measurement Updates

There are essentially just two different "final values" for $e(3|3)$. The larger value of 1.4+ occurs when H_1 or H_2 is used three times in a row. The smaller and value of 1.35 is obtained when some variety exists in choice of H_1 or H_2. It is also noted that the sum (or average) of the three $e(\cdot|i)$ terms has a more selective minimum, and it occurs for the alternating sequences H_2, H_1, H_2 (although H_2, H_1, H_1 gives virtually the same result). This demonstrates that the advantages of alternating measurements discussed in Section 4.3.2 also apply in this case.

The discussion of diminishing returns, in conjunction with Eq. (62) applies approximately to the mean square error terms here. That is,

$$\frac{e(1|1)}{6} \approx \frac{1}{2}, \quad \frac{e(2|2)}{e(1|1)} \approx \frac{2}{3}, \quad \frac{e(3|3)}{e(2|2)} \approx \frac{3}{4}$$

which fits the pattern $\frac{k}{k+1}$ given earlier for the single axis eigenvalues. This hardly constitutes a proof, but it does suggest that in setting up an optimal sequential selection problem (like Eq. (99) and Eq. (100)), some sort of discounting factor should be applied to each successive use of a measurement set with a basic score of S_o. Perhaps $S(k+1) = \frac{kS_o}{k+1}$ would be a reasonable approximation. The determination of S_o would depend upon the geometry, noise levels and state uncertainty existing at the beginning of the sequence. The simple pattern demonstrated here will be disrupted somewhat where $\Phi \neq I$ and $Q \neq 0$. Additional study would be needed to clarify this point.

6. SENSOR SUBSTITUTION STRATEGIES

It is well known that the smallest mean square estimation error results when every available measurement is used. Even a very noisy measurement helps a little, or at least does not degrade accuracy, assuming that the appropriate noise covariance is used. The same statement applies to any secondary sensors such as radar altimeters. If this philosophy was adopted there would be no measurement selection problem. The procedure would be to just use everything.

Hardware and software limitations normally prevent the implementation of the "use everything" approach. For example the number of visible satellites will often exceed the number of receiver channels available for tracking. Therefore choices must be made.

6.1 An Optimization Problem

One approach to these types of problems[14,15,16] is to define an over-all measurement vector consisting of very possible measurement. Each vector subset is multiplied by a scalar m_i

$$\underline{z} = \begin{bmatrix} m_1 \underline{z}_1 \\ m_2 \underline{z}_2 \\ \vdots \\ m_r \underline{z}_r \end{bmatrix} \tag{101}$$

Then, a scalar cost function which depends upon the resulting mean square error (such as Eq. (99) for example) and the cost of taking and processing each measurement is minimized with respect to m_i. A common approach is to use the matrix version of the minimum principle[18], with the elements of the covariance matrix P as state variables. If a given scalar $m_i = 0$, that means the i^{th} group of measurements is not used, while $m_i = 1$ means it is used. Different restrictions on m_i have been considered, including

$\Sigma m_i = 1$[14] and $m_i = 0$ or 1[15,16]. It appears that the essence of the results in[16] are to not include any measurement whose cost exceeds its value, and the "value" looks very similar to what was defined as the score in Eq. (34).

If the optimal selection measurement problem is formulated as in Eq. (101), and if all measurements are included, whether they be from GPS satellites or from miscellaneous alternative devices, then the sensor substitution problem would not need to be considered as a separate issue. Rather, it could be included with topics of Sections 4 and 5.

6.2 A Suboptimal Approach

The approach taken here is a suboptimal one, based on maintaining separate identities of GPS satellite measurements and those derived from nominally less accurate backup sensors.

Let the full state vector $\underline{x}$ which is being estimated be partitioned into two parts, and let the satellite measurements $\underline{z}$ also be partitioned into two parts.

$$\underline{x} = \begin{bmatrix} \underline{x}_1 \\ \\ \underline{x}_2 \end{bmatrix} , \quad \underline{z} = \begin{bmatrix} \underline{z}_1 \\ \\ \underline{z}_2 \end{bmatrix}$$

Then the measurements of Eq. (1) can be written in partitioned form

$$\left[\begin{array}{c} \underline{z}_1 \\ \hline \underline{z}_2 \end{array}\right] = \left[\begin{array}{c|c} H_{11} & H_{12} \\ \hline H_{21} & H_{22} \end{array}\right] \left[\begin{array}{c} \underline{x}_1 \\ \hline \underline{x}_2 \end{array}\right] + \left[\begin{array}{c} \underline{n}_1 \\ \hline \underline{n}_2 \end{array}\right] \tag{102}$$

The question under study assumes that another completely separate method of solving for the $\underline{x}_2$ part of the state is available from alternate (non-GPS satellite) sensors. That is, measurements $\underline{z}_3$ satisfy

$$\underline{z}_3 = H_{32}\, \underline{x}_2 + \underline{n}_3 \tag{103}$$

with H_{32} invertible so that

$$\hat{\underline{x}}_2 = H_{32}^{-1}\, \underline{z}_3 \tag{104}$$

The alternate sensor measurements are assumed inferior to good GPS measurements, [illegible] than degraded or heavily jammed GPS measurements. The [illegible] components which are degraded are contained in the partiti[illegible] o particular restrictions are placed on the size of $\underline{z}_2$. [illegible]ld represent pseudo-range data from one or more satellite[illegible] ...ewise $\underline{x}_2$ could be a single state component (such as altitude from a radar altimeter) or more than one component (such as the horizontal position from an inertial navigator). It is known that the true optimal estimate of $\underline{x}$ would use all three sources of data, $\underline{z}_1$, $\underline{z}_2$ and $\underline{z}_3$ (assuming that the noise statistics for the degraded $\underline{z}_2$ were known). The state update equations from Eq. (12b) can be expressed in the partitioned form as

$$\hat{\underline{x}}_1^+ = \hat{\underline{x}}_1^- + K_{11}(\underline{z}_1 - H_{11}\,\hat{\underline{x}}_1^- - H_{12}\,\hat{\underline{x}}_2^-) + K_{12}(\underline{z}_2 - H_{21}\,\hat{\underline{x}}_1^- - H_{22}\,\hat{\underline{x}}_2^-) + K_{13}(\underline{z}_3 - H_{32}\,\hat{\underline{x}}_2^-) \tag{105}$$

and

$$\hat{\underline{x}}_2^+ = \hat{\underline{x}}_2^- + K_{21}(\underline{z}_1 - H_{11}\,\hat{\underline{x}}_1^- - H_{12}\,\hat{\underline{x}}_2^-) + K_{22}(\underline{z}_2 - H_{21}\,\hat{\underline{x}}_1^- - H_{22}\,\hat{\underline{x}}_2^-) + K_{23}(\underline{z}_3 - H_{32}\,\hat{\underline{x}}_2^-) \tag{106}$$

where superscript - and + indicated before and after the measurement update respectively. One alternate sensor strategy consists of dropping $\underline{z}_2$ entirely and depending on $\underline{z}_3$ to estimate $\underline{x}_2$. That is, a possible algorithm form might be

$$\hat{\underline{x}}_2^+ = H_{32}^{-1}\,\underline{z}_3 \tag{107}$$

$$\hat{\underline{x}}_1^+ = \hat{\underline{x}}_1^- + K'_{11}(\underline{z}_1 - H_{11}\,\hat{\underline{x}}_1^-) + K'_{13}\,\underline{z}_3 \tag{108}$$

Equation (108) reduces to the same form as Eq. (105), if $K_{12} = 0$ and if $K_{13} = -K_{11}\,H_{12}\,H_{32}^{-1}$. Also, Eq. (106) is the same as Eq. (107) if K_{21} and K_{22} approach zero and $K_{23} = H_{32}^{-1}$. Questions of interest are: When can $\underline{z}_2$ be dropped and still maintain near optimal solutions?

Of several potential alternatives (different partitions of $\underline{x}_2$), which one is the best choice? The method of approaching these issues is to first determine detailed expressions for the partitioned suboptimal gain matrix inherent in Eq. (107) and (108),

$$K' = \begin{bmatrix} K'_{11} & 0 & -K'_{11} H_{12} H_{32}^{-1} \\ 0 & 0 & H_{32}^{-1} \end{bmatrix} \tag{109}$$

Then the partitioned form of the optimal gain matrix

$$K = \begin{bmatrix} K_{11} & K_{12} & K_{13} \\ K_{21} & K_{22} & K_{23} \end{bmatrix} \tag{110}$$

is determined. This is done by applying the usual Kalman gain equation

$$K = P H^T [HPH^T + R]^{-1} \tag{111}$$

to the full set of measurements consisting of Eq. (102) plus (103). Then, conditions under which K approaches K' when noise levels $\underline{n}_2$ increase, will be presented.

6.3 The Suboptimal Gain Matrix

The form of the suboptimal gain matrix K' is completely specified once K'_{11} is determined. Motivated by Eq. (103) and the first part of Eq. (102), define an equivalent measurement

$$\begin{aligned} \underline{m} &= \underline{z}_1 - H_{12} H_{32}^{-1} \underline{z}_3 \\ &= H_{11} \underline{x}_1 + (\underline{n}_1 - H_{12} H_{32}^{-1} \underline{n}_3) \end{aligned} \tag{112}$$

If $\underline{n}_1$ and $\underline{n}_3$ are independent and have covariance matrices R_{11} and R_{33} respectively, the equivalent total noise $\underline{v} = \underline{n}_1 - H_{12} H_{32}^{-1} \underline{n}_3$ has covariance $R_{11} + H_{12} H_{32}^{-1} R_{33} H_{32}^{-T} H_{12}^T$. The optimal estimate of $\underline{x}_1$ from Eq. (112) is

$$\hat{\underline{x}}_1^+ = \hat{\underline{x}}_1^- + \mathrm{cov}(\underline{x}_1, \underline{m}) \left[\mathrm{cov}(\underline{m}, \underline{m})\right]^{-1} \left[\underline{m} - E(\underline{m})\right]$$

$$= \hat{\underline{x}}_1^- + K'_{11} \left[\underline{z}_1 - H_{11} \hat{\underline{x}}_1^-\right] - K'_{11} H_{12} H_{32}^{-1} \underline{z}_3 \qquad (113)$$

where $K'_{11} = \mathrm{cov}\left[\underline{x}_1, \underline{m}\right] \left[\mathrm{cov}(\underline{m}, \underline{m})\right]^{-1}$ (114)

Using the covariance definitions

$$\mathrm{cov}\left[\underline{a}, \underline{b}\right] = E\{\left[\underline{a} - E(\underline{a})\right] \left[\underline{b} - E(\underline{b})\right]^T\}$$

immediately leads to

$$K'_{11} = P_{11}^- H_{11}^T \left[H_{11} P_{11}^- H_{11}^T + R_{11} + H_{12} H_{32}^{-1} R_{33} H_{32}^{-T} H_{12}^T\right]^{-1} \qquad (115)$$

6.4 Comparison of the Optimal and Suboptimal Gains

The partitioned form of the a priori covariance (just before the measurement) is

$$P = \begin{bmatrix} P_{11}^- & P_{12}^- \\ (P_{12}^-)^T & P_{22}^- \end{bmatrix}$$

The superscript minus signs will be omitted for convenience in this section. The full measurement matrix is

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \\ 0 & H_{32} \end{bmatrix}$$

and the measurement noise covariance is

$$R = \begin{bmatrix} R_{11} & R_{12} & 0 \\ R_{12}^T & R_{22} & 0 \\ 0 & 0 & R_{33} \end{bmatrix}$$

Substitution into Eq. (111) gives

$$K = \left[\begin{array}{c|c|c} P_{11} H_{11}^T + P_{12} H_{12}^T & P_{11} H_{21}^T + P_{12} H_{22}^T & P_{12} H_{32}^T \\ \hline P_{12}^T H_{11}^T + P_{22} H_{12}^T & P_{12}^T H_{21}^T + P_{22} H_{22}^T & P_{22} H_{32}^T \end{array}\right] U^{-1} \qquad (116)$$

where

$$U = \left[\begin{array}{c|c|c} H_{11}P_{11}H_{11}^T + H_{11}P_{12}H_{12}^T + H_{12}P_{12}^TH_{11}^T + H_{12}P_{22}H_{12}^T + R_{11} & H_{11}P_{11}H_{22}^T + H_{11}P_{12}H_{22}^T + H_{12}P_{12}^TH_{21}^T + H_{12}P_{22}H_{22}^T + R_{12} & H_{11}P_{12}H_{32}^T + H_{12}P_{22}H_{32}^T \\ \hline H_{21}P_{11}H_{11}^T + H_{21}P_{12}H_{12}^T + H_{22}P_{12}^TH_{11}^T + H_{22}P_{22}H_{12}^T + R_{21} & H_{21}P_{11}H_{21}^T + H_{21}P_{12}H_{22}^T + H_{22}P_{12}^TH_{21}^T + H_{22}P_{22}H_{22}^T + R_{22} & H_{21}P_{12}H_{32}^T + H_{22}P_{22}H_{32}^T \\ \hline H_{32}P_{12}^TH_{11}^T + H_{32}P_{22}H_{12}^T & H_{32}P_{12}^TH_{21}^T + H_{32}P_{22}H_{22}^T & H_{32}P_{22}H_{32}^T + R_{33} \end{array}\right] \quad (117)$$

If the submatrix partitions of U are called U_{ij} for convenience, and if, due to R_{22} becoming large, the following terms can be neglected

$$U_{12}U_{22}^{-\frac{1}{2}}\ ,\ U_{32}U_{22}^{-\frac{1}{2}}\ ,\ U_{22}^{-\frac{1}{2}}U_{21}\ ,\ U_{22}^{-\frac{1}{2}}U_{23}\ , \quad (118)$$

$$\text{then } U^{-1} = \left\{\begin{bmatrix} I & 0 & 0 \\ 0 & U_{22}^{\frac{1}{2}} & 0 \\ 0 & 0 & I \end{bmatrix}\begin{bmatrix} U_{11} & U_{12}U_{22}^{-\frac{1}{2}} & U_{13} \\ U_{22}^{-\frac{1}{2}}U_{21} & I & U_{22}^{-\frac{1}{2}}U_{23} \\ U_{31} & U_{32}U_{22}^{-\frac{1}{2}} & U_{33} \end{bmatrix}\begin{bmatrix} I & 0 & 0 \\ 0 & U_{22}^{\frac{1}{2}} & 0 \\ 0 & 0 & I \end{bmatrix}\right\}^{-1}$$

$$= \begin{bmatrix} (U_{11}-U_{13}U_{33}^{-1}U_{23})^{-1} & 0 & -(U_{11}-U_{13}U_{33}^{-1}U_{31})^{-1}U_{13}U_{33}^{-1} \\ 0 & 0 & 0 \\ -U_{33}^{-1}U_{31}(U_{11}-U_{13}U_{33}^{-1}U_{13})^{-1} & 0 & (U_{44}-U_{31}U_{11}^{-1}U_{13})^{-1} \end{bmatrix} \quad (119)$$

If $$P_{12} = [0] \quad (120)$$

then Eq. (116) and (119) yield

$$K = \left[\begin{array}{c|c|c} P_{11}H_{11}^T(U_{11} - U_{13}U_{33}^{-1}U_{23})^{-1} & 0 & -K_{11}U_{13}U_{33}^{-1} \\ \hline P_{22}(H_{12} - H_{32}^T U_{33}^{-1}U_{31})(U_{11} - U_{13}U_{33}^{-1}U_{23})^{-1} & 0 & H_{32}^{-1} \end{array}\right] \quad (121)$$

K_{13} reduces to $-K_{11}H_{12}H_{32}^{-1}$ provided that $U_{13}U_{33}^{-1} = H_{12}H_{32}^{-1}$. K_{21} is zero if $H_{32}^T U_{33}^{-1}U_{31} = H_{12}^T$. Both of these conditions are met if

$$R_{33} << H_{32}P_{22}H_{32}^T \quad (122)$$

Eq. (122), plus a matrix inversion identity, was also used in reducing a long expression for K_{23} down to just H_{32}^{-1}. Finally, the inverse matrix in K_{11} is

$$\begin{aligned} &H_{11}P_{11}H_{11}^T + H_{12}P_{22}H_{12}^T + R_{11} - H_{12}P_{22}H_{32}^T(H_{32}P_{22}H_{32}^T + R_{33})^{-1}H_{32}P_{22}H_{12}^T \\ &= H_{11}P_{11}H_{11}^T + R_{11} + H_{12}\left[P_{22} - P_{22}H_{32}^T(H_{32}P_{22}H_{32}^T + R_{33})^{-1}H_{32}P_{22}\right]H_{12}^T \\ &= H_{11}P_{11}H_{11}^T + R_{11} + H_{12}\left[P_{22}^{-1} + H_{32}^T R_{33}^{-1}H_{32}\right]^{-1}H_{12}^T \end{aligned} \quad (123)$$

Therefore $K_{11} = K'_{11}$ provided that $P_{22}^{-1} << H_{32}^T R_{33}^{-1}H_{32}$. However this is just an alternate form of Eq. (122). Therefore $K \rightarrow K'$ provided R_{22} becomes sufficiently large so that terms (118) are negligible, and Eq. (120) and (122) are satisfied.

6.5 Application of the Substitution Strategy

If a given measurement or group of measurements $\underline{z}_2$ becomes very degraded, so that R_{22} increases to a very large value, then an alternate method of solving for one or more states $\underline{x}_2$ should be considered. If more than one substitution choice is available, the one which gives $\underline{x}_2$ most nearly uncorrelated from $\underline{x}_1$, should be selected so that Eq. (120) is most nearly true. The substitution should be made when P_{22} becomes sufficiently degraded (due to the increasing R_{22}) so that Eq. (122) is reasonably satisfied.

That is, under normal operation it is assumed that the secondary sensor is less accurate than GPS navigation accuracies, so $R_{33} > H_{32} P_{22} H_{32}^T$. Under degraded conditions the sense of the inequality will eventually reverse.

Some of the ideas presented above are similar to those given previously, for other purposes, by Stubberud and Wismer[19]. The requirement that the state vector be partitioned into the most nearly uncorrelated subvectors $\underline{x}_1$ and $\underline{x}_2$ is one of the major points of similarity. The best method of verifying the appropriateness and effectiveness of this procedure is through realistic simulation tests.

7. ANALYTICAL TESTS AND SUGGESTED SIMULATION CASES

7.1 Analytical Tests

As stated in Section 3, one of the tasks of this study was to carry out analytical tests of the proposed GDOP generalizations and extensions. This task has been accomplished in earlier sections concurrently with methodology development. Specifically, in Section 4 examples were presented which demonstrated that the best GDOP does not guarantee highest accuracy. It was shown that the eigenvalues and eigenvectors of the covariance matrix play a very important role, in addition to geometry and noise levels. Analytical tests were given in low dimensional cases, and these were used to extract some general principles and rules of thumb for higher dimensional cases.

In Section 5 analytical tests on a low dimensional model demonstrated that the measurement types which are best for the instantaneously minimum position error need not give the minimum average position error over time. A law of diminishing returns for repeated use of the same measurement was developed analytically, as was the suggestion that a sequence of measurements will be more effective if there is some variety in the sequence.

7.2 Suggested Simulation Cases

A considerable amount of computational work was carried out in conjunction with this study. Some of it applied to simple examples, but much of it applied to realistic GPS geometry. The measurement noise model used in these cases was a hypothetical one. The a priori state covariance matrices were selected parametrically. It is suggested that the methods of Section 4, and in particular Section 4.4, be tested in a realistic full blown simulation with the latest noise models and covariance matrices. Even the geometry, which was chosen realistically at the time of this study, deserves additional attention because of the smaller number of satellites

in the scaled-down GPS network. This affects viewing times and reduces the number of potential measurement choices at any given time. One type of case that is expected to benefit most from generalized GDOP methods has a highly eccentric state covariance error ellipsoid, a wide range of measurement noise levels, and GDOP values that are all reasonable or roughly comparable with each other.

In Section 4.5 it was shown that root mean square position error could be improved by a few feet up to maybe 10 feet at a given measurement time by using a generalized GDOP. Although small, these improvements are significant in view of GPS accuracy goals. This level of static improvement should be verified with a full and realistic simulation. Such a simulation could also determine how these relatively small improvements in accuracy propagate in time. Do they remain small or do they accumulate, and if so, how much?

The entire subject of sequential measurements needs to be studied by means of simulation. One scalarized version of the problem was formulated in Section 5.2. Some remaining details of that method can only be extracted from actual simulations. If that scalar approach to setting up the problem proves feasible, then a discrete time dynamic programming solution over a finite planning horizon can probably be obtained in near real time. If the scalar approach is not satisfactory, then the higher dimensional full matrix approaches of References [13] through [16] could be considered. The additional complexity would be a disadvantage for real time applications.

A reasonably simple sensor substitution strategy was detailed in Section 6. This should also be evaluated by simulating one or more very bad pseudo-range measurement channels. The partitioning of the state into essentially uncorrelated subparts is a general requirement that can only be tried in simulation. Also, several

matrix inequality conditions were developed in Section 6.4. Exactly what is meant by one matrix being small compared to another was not stated. Several measures of relative smallness can be considered, but only simulation can give a feeling for which one works best.

8. CONCLUSIONS AND RECOMMENDATIONS

This study has shown that measurements selected because they have the best GDOP do not necessarily give the smallest estimation error. In fact examples have been presented in which the worst GDOP measurements give the best position estimate.

Generalizations to GDOP have been developed. Some of these are approximations, and one is exact. Performance improves, but at the expense of requiring more input data and a higher computational burden. The ability to minimize a weighted sum of the estimation error components is built into the exact algorithm and this may be useful in some applications. In addition to the development of an exact algorithm for ranking measurement effectiveness, eigenvalue-eigenvector theory was used to obtain insight into the tradeoffs among geometry, noise levels and a priori uncertainty.

When a sequence of measurements are taken over a period of time, instantaneously smallest position errors may not give the smallest time average position errors. There is a position-velocity error tradeoff that relates to the error growth between measurements as well the error reduction at a measurement. This question is dealt with in detail for a simplified model, and a proposed approach is given for a more realistic model.

When one or more satellite channels are lost or heavily jammed, the recursive Kalman filter can carry on with a smaller measurement vector. This will degrade system accuracy, so it may be advantageous to consider substituting backup sensors in place of the missing or degraded satellite signals. A suboptimal partitioned state vector approach to the problem has been developed, along with some guidance regarding which substitute sensor ought to be selected, and under what conditions the substitution ought to be made. The answers to these questions depend upon the

correlations that exist among state components, and upon the relative accuracy levels of the current state estimate, the failed sensors and the substitute sensors.

The duration and scope of this study did not allow resolution of all details of the many issues which were raised. To do this will require a concentrated effort with a full blown simulation, including realistic geometries, noise models, operational restrictions and data processing limitations. Three distinct studies of this variety can be imagined: the optimal static measurement selection problem, the dynamic or sequential measurement selection problem, and the sensor-substitution problem. Suggested directions for further study and simulation tests have been given for each of these areas.

REFERENCES

1. R. J. Milliken and C. J. Zoller, "Principle of Operation of NAVSTAR and System Characteristics," Navigation: Jour. of Institute of Navigation, Vol. 25, No. 2, Summer 1978, pp. 95-106.

2. J. J. Spilker, "GPS Signal Structure and Performance Characteristics," Navigation: Jour. of Institute of Navigation, Vol. 25, No. 2, Summer 1978, pp. 121-146.

3. W. L. Brogan and A. C. Liang, "Navigation of a Tactical Aircraft Using a Fully Integrated DNSS/Strapped-Down Inertial System," Aerospace Report No. TOR-0073(3020-03)-2, March 4, 1973, pp. 28-30.

4. P. B. Liebelt, An Introduction to Optimal Estimation, Addison-Wesley, 1969.

5. J. S. Meditch, Stochastic Optimal Linear Estimation and Control, McGraw-Hill, 1969.

6. W. L. Brogan, Modern Control Theory, Quantum Publishers, Inc., New York, 1974.

7. P. M. DeRusso, R. J. Roy and C. M. Close, State Variables for Engineers, John Wiley and Sons, Inc., New York, page 234.

8. R. K. Mehra, "Optimization of Measurement Schedules and Sensor Designs for Linear Dynamic Systems," IEEE Trans. on Automatic Control, Vol. AC-21, No. 1, February 1976, pp. 55-64.

9. W. L. Brogan, "Improvements and Extensions of the Geometrical Dilution of Precision Concept for Selecting Navigation Measurements," IEEE Position, Location and Navigation Symposium, Atlantic City, December 1980.

10. E. Dudzinski, Interim Report WU 60951410: "GPS Satellite Selection," Air Force Avionics LAB/RWA, 14 January 1980.

11. K. C. Bell, P. L. Konop and B. A. Kriegsman, "HPAA Performance Evaluation in a Multiple Jammer Environment," Report C-5220, ADGINT Program, C.S. Draper Labs, May 1979.

12. W. L. Brogan, "Second Quarterly Progress Report on Improvements and Extensions of the Geometrical Dilution of Precision Concept for Selecting Navigation Measurements," Letter Report on USAF Contract F33615-80-C-1092, November 15, 1980.

13. L. Meier, III, J. Peschon and R. M. Dressler, "Optimal Control of Measurement Subsystems," IEEE Trans. on Automatic Control, Vol. AC-12, No. 5, 1967, pp. 528-536.

14. M. Athans, "On the Determination of Optimal Costly Measurement Strategies for Linear Stochastic Systems," Automatica, Vol. 18, July 1972, pp. 397-412.

15. K. D. Herring, "Measurement Subsystem Parameter Optimization for Estimation Applications," Ph.D. Dissertation, Information and Control Sciences Center, Southern Methodist University, December 1972.

16. K. D. Herring and J. L. Melsa, "Optimum Measurements for Estimation," Preprint of paper based on Ref. 15, circa 1972.

17. N. J. Mejstrik and W. L. Brogan, "A Stochastic Optimal Path Selection Problem," International Journal of Control, Vol. 15, No. 6, pp. 1063-1072, 1972.

18. M. Athans, "The Matrix Minimum Principle," Information and Control, Vol. 11, 1968, pp. 592-600.

19. A. R. Stubberud and D. A. Wismer, "Suboptimal Kalman Filter Techniques," Chapt. 5 in Theory and Application of Kalman Filtering, C. T. Leondes, Editor, AGARDograph 139, February 1970.